

The importance of coherent defence strategies

Astronomers served by the European Space Agency run the risk of undermining their own future by undue criticism of it, and by underestimating the obstacles to imitating NASA.

Europe's particle physicists can face the future with a fair degree of satisfaction. Following recent international endorsements of CERN, they are set to occupy centre stage in the next generation of facilities at the high-energy frontiers of their discipline. European astronomers, in contrast, are rightly worried about their future. In particular, as a result of the downward pressure on budgets, those depending on space platforms find themselves in a new and insecure world where they can no longer rely on the European Space Agency (ESA) to provide them with a nourishing torrent of data. But whether, as some claim, their future can be much invigorated by following the example of the US National Aeronautics and Space Administration (NASA) is doubtful.

ESA's Horizons 2000 programme has provided a relatively secure environment for astronomers, combining stability, through regular launches of large cornerstone missions, with flexibility, through intermittent medium-sized missions which are open to competition from the research community over significantly shorter timescales. The ability of ESA to plan reliably for the future has long been envied by US astronomers. But ESA's member states are no longer able to pay for the visions approved in the 1980s, and rethinking is required. The first report to emerge, led by ESA's Space Science Advisory Committee (SSAC), proposes maintaining the Horizons 2000 missions more or less as they are, but stretching out their launch schedules.

A prospect of periods of data starvation is causing agitation among astronomers in Europe. And that may be the reaction that SSAC is hoping for, acting as it is from a genuine sense of gloom but also, no doubt, in a spirit of political agitation. If the SSAC is hoping that such worries will cut much ice with research ministers from ESA's member states, it is deluding itself. More productively, however, its preliminary recommendations will concentrate ESA's mind on the necessity to question all traditional assumptions in order ruthlessly to cut costs.

Even more pressure in that direction comes from Germany, France and Britain, whose governments, and some of whose astronomers, are calling aggressively for ESA to take lessons from NASA's streamlining of its programmes according to the slogan "smaller, faster, cheaper, better". It is the obvious route to healthy science and healthy balance sheets, say ESA's critics, and if NASA can do it, so can Europe.

Such words are cheap but can also be powerful, and they need to be used carefully when governments have their budget knives sharpened. NASA's approach has not yet been tested — its programme has been held up by launcher problems. Furthermore, the need to develop technologies to deliver efficient, fast, cost-effective science puts new costs up front — NASA's ability to develop them cost effectively has yet to be demonstrated. But NASA has benefited greatly from its inheritance of declassified technologies in cryogenics, detectors, adaptive optics and the like, to which ESA does not have access. And the strong radical leadership possible at NASA cannot readily emerge within ESA, bound up as it is in the inescapable Euro-paraphernalia of negotiations around member-state interests.

To its credit, ESA is reducing several components of its inertia, albeit belatedly. Its director of science, Roger Bonnet, has set up task forces that will report to the SSAC at the end of this month, to investigate in detail how NASA is introducing cost-savings and shortening its time to launch, and deciding which ideas can be imported. Individual missions and organizational procedures are again being scrutinized for cost savings. For these reasons, the SSAC's next report will no doubt be more trenchant about changes to be made within ESA, ahead of the crucial decisions about the Horizons programme's future to be made later this year. In the meantime, critics of ESA who nevertheless wish it well should choose their words carefully. CERN's success at surviving member-state pressures has been as much a result of the coherence among its research community as of anything else. □

Will Japan's money for science be short lived?

Japan's scientists should take note of criticism of the Japan Key Technology Centre.

Japanese scientists, long seen as poor cousins of their Western counterparts, have been enjoying an unprecedented bonanza of extra government funding in recent years. And, if the five-year plan mapped out by the government last year is to be believed, the money should continue to flow liberally for quite some time.

But the first signs of a backlash against the extra money for science may be appearing. In recent weeks, Japan Key-TEC, an organization that was set up with income derived from privatization of Nippon Telegraph and Telephone Corporation, has come under criticism for failing to deliver a significant return on more than ¥200 billion in investment over the past ten years (see page 424).

Proponents of Japan Key-TEC argue that it is unreasonable to expect an immediate return on the basic research that has been abundantly funded by the centre at institutes such as the Biomolecular

Engineering Research Institute (BERI) in Osaka. That is true. But BERI is an exception among Japan Key-TEC institutes. Most of the centre's 60 or more institutes have made little or no impact on the world of science or technology and are being quietly closed down.

Japan's government and university scientists are likely to face similar criticism a few years hence if they fail to produce visible results. The science community must act now to ensure that, through rigorous peer review, the extra money goes to deserving scientists and that particular individuals in priority fields do not get a surfeit of funds. Furthermore, greater efforts must be made to publicize the achievements of institutions and individuals through, on the one hand, objective external reviews and assessment and, on the other, media releases that explain the science to the general public. Otherwise, Japan's scientists may soon become poor cousins again. □



Planning rejection leaves British nuclear waste plans in disarray

[LONDON] Britain's radioactive waste disposal plans were in disarray last week as Nirex, the company charged with the job of waste disposal, effectively abandoned attempts to build an underground repository at Sellafield in the northwest of England.

At a meeting of Nirex's board last week, members decided not to appeal against the British government's decision three weeks ago to refuse Nirex's application to build an experimental research laboratory near its proposed site for a waste repository at Longland's Farm, in Cumbria.

Nirex was hoping to use the laboratory to investigate the Sellafield geology. Its decision not to appeal effectively spells the end of Sellafield as a proposed repository site. The company will now have to turn elsewhere in its quest to dispose of waste from nuclear power stations.

The company also announced "adjustments to [its] programme and resources", widely understood to mean staff redundancies. Contracts for the research laboratory have already been cancelled. And one well-placed source suggested that Nirex's rock laboratory department would almost certainly be shut down, with the loss of scientists.

A clearly disappointed Nirex chairman, Sir Richard Morris, complained in a statement that the ground rules had kept changing. What should have been a straightforward planning inquiry into a research facility turned into a full-scale investigation of the merits of disposing of nuclear waste at Sellafield.

"We cannot get the information to show whether the site is safe or not without a rock laboratory, but it now appears we cannot win approval for a rock laboratory without first showing the site is safe."

But some of Nirex's own internal ground rules are also now coming under scrutiny, particularly the methods employed by the company in its attempts to obtain and then assess the safety of Sellafield as its chosen repository site.

The aim of the laboratory was to test the geology of the rocks, in particular to determine the risk of radioactive waste finding its

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

No future: an aerial view of Nirex's test borehole at Longlands Farm near the Sellafield plant.

way back to the surface once the containers had corroded, and the waste had leaked out.

Cumbria County Council, the local authority, rejected planning permission for the laboratory, backed by environmentalist groups such as Friends of the Earth. FoE argued that surface-based investigations remained incomplete.

A leaked internal memo from John Holmes, Nirex director for science, concurred with this view, and suggested the need for more data sampling points (see *Nature* 385, 282; 1997).

Nirex appealed against the council's refusal, and the matter was passed to Britain's environment secretary John Gummer following a 66-day public inquiry. The inspector presiding over the inquiry recommended that Gummer, too, should refuse the application.

The leaked memo is believed to have played a crucial role in this decision, as Holmes appeared to acknowledge that, unless lower values for Sellafield's rock permeability were found, "we may struggle to

make a case for the site". Gummer said in his rejection that he was "concerned about the scientific uncertainties and technical deficiencies in the proposals presented by Nirex".

The refusal is a huge blow to the company, which has already spent more than £200 million (US\$320 million) assessing the Sellafield site. Nirex was so confident that the government would allow further site investigations that it even awarded initial contracts to build the underground laboratory before Gummer's decision was announced.

The decision lays wide open the question of what to do with the estimated 300,000 cubic metres of radioactive waste that is expected to accumulate by early next century. Nirex was expecting to begin placing waste in a repository by this time. One of its shareholders, British Nuclear Fuels, is believed to hold waste disposal contracts from other countries.

FoE has always insisted on 'retrievability' being built into any disposal plans and ideally would like Nirex to consider disposal at a shallow site. But with deep disposal having become the standard method internationally of disposing of nuclear waste, a shallow repository is unlikely.

The Labour party, which is widely expected to form the next government after the 1 May election, is believed to support deep disposal. The government was poised to publish its plans for a research strategy for disposing of high-level waste, before the election was called. Some observers now speculate that a deep and retrievable repository for intermediate level waste as well as high-level waste may now be the best way forward.

In the meantime, however, Gummer's decision also raises critical questions about whether Nirex itself is equipped to continue with the responsibility of looking for a suitable site, or whether this job should fall to a new body less beholden to the nuclear industry.

The selection of Sellafield as the chosen site has long been claimed by critics to have been made on grounds that were suited more to the needs of the industry than of science. Notwithstanding geological issues, a Sellafield-based repository offered clear advantages over other sites as much of Britain's nuclear industry, including British Nuclear Fuels, is already based there.

John Gummer ended his statement on refusing Nirex's planning appeal by saying that he too was "concerned about the process of site selection". His predecessors are widely believed not to have opposed Sellafield as a potential repository site.

An acute lack of transparency has been a

Nature monthlies on the move

The US editorial offices of *Nature Genetics*, *Nature Medicine* and *Nature Structural Biology* will shortly move from Washington DC to join *Nature Biotechnology* in New York. The US editorial office of *Nature* itself will remain in Washington.

From 1 May 1997, editorial enquiries, manuscript submissions, reviews and proofs for all monthly journals should be sent to 345 Park Avenue South, 10th floor, New York, NY 10010-1707. More information may be found in a full-page advertisement following the Book Reviews section in this issue.

The Rock Lab would have been ideal for burying all this material...
With room for all that nuclear waste as well...



recurring complaint of critics of Nirex for several years. Several external bodies, including the Royal Society, and more recently the government's Radioactive Waste Management Advisory Committee (RWMAC), had called on Nirex to open up its research to independent peer review.

Another criticism is that scientists at Nirex were under pressure to deliver results by unrealistically short deadlines set by senior management. "There is never any time for reflection or analysis," says one former Nirex employee. "We just drill borehole after borehole." (Despite requests, no response had been received from Nirex on these points as *Nature* went to press).

Rachel Western, senior nuclear officer for FoE, says the public has lost trust in Nirex, and that the government should have taken a tighter regulatory line.

She says it is time for the creation of a new and independent body that meets the original terms of a 1976 Royal Commission on Environmental Pollution report on radioactive waste management, known as the 'Flowers report', which first called for the creation of a waste disposal agency. "This report made clear that such an agency had to be independent of the nuclear industry."

But others, such as Sir John Knill, former chairman of RWMAC, disagree on the grounds that representation from the nuclear industry is important if the agency is to have influence within the industry. "I do not believe that the body need be totally independent and the Flowers report does not suggest this," he says.

Morris of Nirex says that the company is now "happy to discuss with government and all other interested parties the future direction of government policy and how it can be implemented with public confidence".

Knill says that "an effective balance could be struck between nuclear and non-nuclear interests ensuring that particular expertise was properly represented". But he emphasizes that "the body would have to operate in a transparent manner and be responsive to the need to gain public confidence". **Ehsan Masood**

Japanese technology fund faces ministry criticism

[TOKYO] The Japanese government is to review the activities of the Japan Key Technology Centre (Japan Key-TEC), which funds more than 60 semi-private research institutes, after public criticism of the effectiveness with which the research funds are being used.

A total of ¥220 billion (US\$1.78 billion) has been spent by the organization since 1985. But the Ministry of Finance is said to be among those concerned that this investment has produced very little return, and that the return is unlikely to increase significantly in the future. Some see the ministry as attempting to raise a question mark over the government's promise to make a massive increase in research spending.

Although the government's budget for the fiscal year 1997 (which began on 1 April) is the tightest for nine years, it will increase spending on science and technology by 11.7 per cent (see *Nature* 385, 104; 1997). Included is an increase of ¥20 billion for research at Japan Key-TEC.

The Ministry of Trade and Industry (MITI), which oversees Japan Key-TEC with the Ministry of Posts and Telecommunications (MPT), announced last week that it intends to take steps to increase the productivity of the research projects it funds. One strategy will be to increase support for small to medium-sized companies, and to review the activities of research institutes considered to be unproductive.

Japan Key-TEC was set up by MITI and MPT in 1985, following the privatization of Nippon Telegraph and Telephone Corporation, to distribute the income received from government-owned shares in the company (see *Nature* 359, 577; 1992).

The organization's goal is to promote collaboration between government, industry and academic institutions to enhance research in areas such as biotechnology and telecommunications. Up to ¥26 billion a year has been spent on institutes that are jointly set up by two or more private companies, which contribute 30 per cent to the initial capital investment, with Japan Key-TEC providing the rest.

Some of these institutes have built substantial reputations in basic research. One is the Biomolecular Engineering Research Institute, formerly known as the Protein Engineering Research Institute (PERI), which is supported by 18 companies including Takeda Chemical Industries, Hitachi, Toray Industries and Japan Tobacco.

Each institute receives funding for between five and ten years to complete an agreed research programme. Of 68 institutes

set up since 1985, 43 have completed their research programmes and now consist of only a handful of administrative staff retained to handle accounts and analyse research data.

One or two of these institutes, such as PERI, have been re-established under a different name and given a new lease of life. But most of the institutes have fixed terms, to prevent research funding from getting locked into particular fields. The research programmes are expected to file patents, royalties on which are eventually intended to be returned to Japan Key-TEC.

By March last year, 3,187 applications for patents had been made, of which 435 had been accepted and 22 had found successful commercial applications. But, according to MITI, the total return of royalties has been only ¥1.3 billion — a small fraction of the total initial investment of ¥220 billion.

In 1992, the government's Management and Coordination Agency asked the ministries to submit research reports from institutes that had completed their programmes, to recover investment from profitable companies and to close down unprofitable ones.

In reply, both MITI and MPT insisted that returns would come in the long term, and showed little intention of liquidating unprofitable institutes. An official from MITI says that "it is possible to yield profit as long as it is regarded as a long-term project" and that "results from basic research should not be accounted solely in financial terms".

Recent criticisms of Japan Key-TEC projects may be the first sign of a backlash against the government's 'five-year plan', announced last June, to increase spending on science and technology by 50 per cent over the next five years.

The government as a whole has been supportive of increased funding for research and development. But, because of Japan's economic difficulties, the Ministry of Finance may find the forecast expenditure too costly.

According to the reform plan issued last week by MITI, companies that have completed their term and those that are unlikely to yield any profit may be closed down, starting as early as next year. But the director of one of the research institutes thinks this is "highly unlikely" to happen.

There are also plans to increase capital investment in small and medium-sized venture companies — a big change from the previous policy where investment was made only in new research and development institutes. The new plan would bring dividends and profit on the sale of shares after the company's flotation.

Asako Saegusa

Pay rises eat into South African budget

[CAPE TOWN] At first glance, research appears to have fared well in the first budget of South Africa's new Finance Minister, Trevor Manuel, presented last month. Total allocations to the research councils increased by 15 per cent — from R993 million (US\$224 million) to R1,128 million — well in excess of both the 6 per cent general budgetary increase and the country's inflation rate of 9 per cent.

But university researchers are unhappy that the Foundation for Research Development (FRD), the main funding agency for research, was awarded an increase of only 4 per cent in the funds it distributes to universities. Most of this money will go to salary increases to the staff employed by the research councils, whose research output, as measured by international citations, is notoriously low (see *Nature* 384, 11; 1997).

Some see an irony in the fact that the FRD was given top rating of all the research councils by the advisory committee that made recommendations to the government on how the science budget should be distributed. The committee's hearings were held in public for the first time last October, and its recommendations were published.

Rob Adam, deputy director-general at the Department of Arts, Culture, Science and Technology, says that the committee's recommendations were followed stringently. But he says that salary increases, over which the department has little influence, overshadowed reallocations made on the grounds of merit.

Adam adds that the reallocation of funds this year was "deliberately conservative", as the research councils are under review. There was little to be gained by making major reallocations before the process of rationalization that is expected to result from the review.

Dave Woods, vice-chancellor of Rhodes University, says that the budget will have the effect "of improving salaries of researchers employed by the councils at the expense of fostering competitive research". He adds that

Allocations to South Africa's science councils	
% increase 1996-97 to 1997-98	
Council for Scientific and Industrial Research	19
Council for Minerals Technology	14
Council for Geosciences	8
Agricultural Research Council	11
Medical Research Council	14
Human Sciences Research Council	11
Foundation for Research Development	4
National Accelerator Centre	9
S. A. Astronomical Observatory	10
Hartebeespoort Radio-astronomical Observatory	9
Crime prevention	10
Special projects	52
Total	15

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

South Africa's Antarctic base took so long to finish that it sports the colours of the old national flag.

university salaries are lagging behind those in government and the research councils.

Total funding for tertiary education was increased by 11 per cent in the budget, with universities receiving 66 per cent of their entitlement under South Africa's complex funding formula (see *Nature* 384, 13; 1997). Although this is less than last year, when they were awarded 68 per cent, it is more than the 60 per cent the government had proposed earlier.

Sibusiso Bengu, Minister of Education, has denied that this increase was agreed in response to widespread campus protests which accompanied the start of the academic year in February.

Another anomaly in the budget is the limited funding allocated to research in Antarctica. In January, South Africa opened a new base there at a cost of R80 million. But only R14 million has been allocated to the Antarctic research programme, of which R11 million will be spent on supplying and running the base. That leaves just R2.6 million for the Antarctic research programmes in geology, biology, oceanography and physics.

"We could do a lot more with a little more money," says Dave Walker of the University of Natal, who heads the physics programme that forms part of the Southern Hemisphere Auroral Radar Experiment. **Michael Cherry**

Developing world wired up to molecular data

[LONDON] The Weizmann Institute in Israel launched on 31 March an expanded communications network to provide countries in eastern Europe and Asia with access to on-line molecular databanks and training in their use.

The International Center for Cooperation in Bioinformatics (ICCB), which is sponsored by the United Nations Educational, Scientific and Cultural Organization (Unesco), will provide on-line data on genes and proteins, according to Luba Vikhanski, a spokesman for the Weizmann Institute.

"Computerized resources in molecular biology are now crucial for the progress of medicine, biotechnology and agriculture," says Vikhanski. Scientists in the United States, Japan and Western Europe have access to their own data networks, he says, but "their colleagues in most other regions lag far behind".

Scientists in Western Europe, for example, are connected to the European

Molecular Biology Network, EMBnet, which is based in the United Kingdom. Israel is connected to EMBnet and the Weizmann Institute hosts the Israeli node, providing access and training in bioinformatics to 250 research groups in Israel. "The ICCB will be the equivalent of EMBnet for the developing world," says Vikhanski.

Over the past three years, the institute has set up Unesco-sponsored activities in bioinformatics in Poland, Turkey and India, and most recently in China. These countries will now form the regional headquarters of the ICCB network, providing network access to their neighbours. Israel will serve as the central node, in addition to providing network access to the Middle East region.

Much of the discussion at a two-day inaugural meeting, which ended on 1 April, focused on the centre's future funding. The network cost US\$256,000 to set up, \$200,000 of which was provided by Unesco. The rest was provided by Israel's foreign ministry, the Weizmann Institute and industry. □

France urged to head 'Rubbiatron' efforts

[PARIS] France's ministry of research has been urged to head an international collaboration to develop ideas drawn up by the physicist Carlo Rubbia for a 'safe' nuclear reactor — and in particular to examine using the technology to destroy nuclear waste.

The proposal comes from Claude Birraux, a member of the National Assembly who has represented the Haute Savoie region since 1978. Birraux, a chemical engineer by training, has been the author of a number of reports on nuclear safety produced by the parliament's office of science and technology assessment.

Last November, Rubbia — Nobel prize-winner in 1984 for discovering the W and Z particles — presented members of the parliamentary office with proposals for a coupled accelerator-reactor using thorium as fuel. It was dubbed 'Rubbiatron' by the French press.

Rubbia has already received funding from the European Commission for research into his ideas at CERN, the European Laboratory for Particle Physics in Geneva, Switzerland, of which he was formerly director.

The proposed reactor would use an accelerator producing a current of protons, which would bombard molten lead to produce high-energy neutrons. These would be directed at thorium-232 targets, which transform into fissile uranium-233, starting a chain reaction.

According to Rubbia, the system has several advantages over conventional reactor designs. The reactor would remain subcritical, and it could be switched off rapidly by cutting the proton beam. Also, replacing thorium with either plutonium or uranium-239 would allow the reactor to burn nuclear waste.

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Ring of promise: Experiments on Carlo Rubbia's proposed reactor ideas at CERN in Geneva.

The production of radioactive actinides would be much lower than with conventional reactors, says Rubbia. The cost of energy production would also be competitive, although the initial costs would be considerable. A prototype 100-MW reactor would cost FF1 billion (US\$175 million).

But Birraux's report on Rubbia's ideas, published in draft form last month, highlights many potential problems. Many physicists feel that there remain major technical problems with the design of the tungsten 'door' — the 'admission window' — by which protons would enter the reactor. This would have to resist high-energy protons and neutrons as well as the corrosive power of molten lead.

The window would have to be replaced once a year, causing difficulties that will need to be taken into account in its design. And there are major uncertainties over the thermohydraulic properties of molten lead, as

well as its highly corrosive behaviour.

Birraux says these problems will take time to resolve. "I don't believe, like Rubbia, that a prototype could be achieved within five to six years," he says. But he sees considerable potential in the machine as a way of disposing of nuclear waste. A law passed in 1991 requires research to be carried out on nuclear safety and the burning of nuclear waste, and Birraux argues that this could cover research on the Rubbiatron.

The French government is not due to make a final decision on a waste-disposal technique until 2006, and Birraux considers this gives sufficient time to carry out experiments based on Rubbia's ideas. France currently lacks a proton accelerator, and one built for use in Rubbia's project would have many other uses even if the project itself fails. The Phénix fast reactor is planned to be decommissioned by 2005, after which France will be left without a high-speed neutron source.

Birraux's report therefore recommends that the accelerator should be built, while research should continue on the properties of molten lead, the admission window and thorium cycle. It insists on a practical strategy with clear and specific goals. "We must avoid a second Superphénix story," says Birraux.

Rather than proposing that the French government should contribute directly to Rubbia's efforts, Birraux calls for a wide international collaboration. A Euratom commission is due to discuss whether funding for the project should be included in the European Union's fifth Framework research programme. That commission has produced a report identifying the use of Rubbia's technique to burn actinides as a top priority.

Birraux points to knowledge gained from Russia's experience in using molten lead as a coolant in nuclear submarines, and that by the Paul Scherrer Institute in Switzerland from its high-current accelerators.

A key issue remains who would oversee a collaborative project. France's Atomic Energy Commission already has its Isaac programme studying a coupled accelerator-reactor. Even though this is in its early stages, the organization's experience in nuclear issues would be required. CERN is prohibited from carrying out research related to nuclear energy. Birraux suggests that the research ministry could act as a mediator in helping to encourage cooperation.

Cooperation is certainly important, given the cost involved and the divergence between Rubbia's enthusiasm for a reactor and French interest in more immediately practical goals. Birraux is not totally convinced of Rubbiatron's capacity to provide electricity at competitive rates, and feels that Rubbia may be underestimating the technical difficulties. "It's too soon to know," he says. **Eric Glover**

Closure looms for Hungarian institutes

[MUNICH] The Hungarian Academy of Sciences will close seven of its 44 research institutes and shed up to a quarter of its staff if a proposal approved by its senate last week is endorsed by the general assembly in May.

In common with almost all scientific organizations in eastern Europe, the academy has been struggling to survive rapidly falling budgets since the reorganization in 1989 that followed the fall of Communism (see *Nature* 367, 503; 1994).

László Keviczky, secretary general to the academy, is reluctant to name the institutes put forward for closure or merger, but those specializing in physics and chemistry are expected to be worst hit. The academy's life sciences and mathematical sections are not expected to be seriously affected.

The closure of seven institutes and the slimming down of others is expected to result in the loss of about a thousand jobs.

The organization decided last year that it

should no longer try to maintain its current strength. Its 400-strong general assembly set up a special commission to draw up proposals — which the senate has now approved — for changes needed to keep the academy viable.

"Reductions are necessary to preserve the scientific excellence of the academy's institutes," says Keviczky. "This means change to the quantity as well as the quality of research."

But Keviczky is worried that the general assembly may not be enthusiastic about the closure of individual institutes, despite its formal recognition of the academy's need to contract. "It will take a very strong effort to make this historic decision," he says.

Academy staff numbers have already fallen by more than a third since 1990 as scientists have left the country or joined industry, mostly because of low wages and poor working conditions. **Quirin Schiermeier**

New challenge to nuclear test facility

[WASHINGTON] Environmental groups are expected to mount a last-ditch court action this week to block construction of the National Ignition Facility (NIF), which is due to start at Lawrence Livermore National Laboratory in California.

A total of 50 groups, led by the Natural Resources Defense Council (NRDC), will ask the Federal District Court in Washington for a preliminary injunction to block construction work. They will argue that the Department of Energy (DoE) failed properly to consider alternatives to its Science-based Stockpile Stewardship Program, of which the NIF is a central component, for maintaining the US nuclear weapons stockpile.

The environmentalists will charge that an extensive Programmatic Environmental Impact Statement, completed by DoE last November, failed to assess broad alternatives such as the remanufacture of weapons to existing designs.

The NIF will attempt to ignite tiny pellets of deuterium-tritium fuel by firing 192 extremely powerful laser beams at it. Ignition occurs if fusion in the fuel releases enough energy to sustain itself. A study by the National Academy of Sciences, released last month, concluded that current scientific

understanding gives "a reasonable expectation that ignition will be achieved".

But both opponents and proponents of the NIF argue that DoE is keen to build it — whether it achieves ignition or not — because of the weapons-physics experiments it will do and because of its role in retaining top-rate physicists at weapons laboratories.

The academy report lists a number of targets set in 1990 for Nova, the existing inertial confinement fusion machine at Livermore, which have not been met. These include maximum temperature at the target and the "convergence ratio", a measure of how tightly pellets are compressed.

But the report still concludes that the NIF will make "important contributions" to stockpile stewardship, and should proceed as planned. Although court action by the NRDC blocked the DoE from using the academy report, the department decided to proceed with the project in any case (see *Nature* 386, 209; 1997).

Steve Koonin of the California Institute of Technology, who chaired the academy panel, is confident that the problems can be overcome. But he said it was impossible to put a number on the "reasonable" probability that the machine would achieve

ignition: "Reasonable means to me that a prudent physicist would go ahead and do it."

Koonin said that he was proud to have participated in the academy process, and criticized Tom Cochran, the senior arms control analyst at the NRDC. "I only wish that Tom Cochran was subject to the same kind of peer review that the committee was subject to." He said that "there wasn't much technical expertise" in the NRDC's submissions to the panel.

Cochran was unavailable for comment but Christopher Paine, another NRDC analyst, accused the academy committee of bias towards NIF. "We don't say these people are dishonorable — just that the committee was unbalanced", he said. Paine pointed out that the committee estimated that NIF would now cost \$1.6 billion, including Livermore's own costs, against the widely assumed figure of \$1.1 billion.

"The days of building \$1-billion machines that were fantastic flops are gone," he said. "The reason NIF has got so far is that it is riding on the back of stockpile stewardship. The preponderance of evidence is that it isn't ready — the academy concedes as much when it says it has a 'reasonable' chance of working."

Colin MacLachlan

White House science office 'needs more focus', says panel

[WASHINGTON] The Office of Science and Technology Policy (OSTP) in the US White House should have a more tightly defined role in the process of setting the president's annual budget and should employ a chief of staff, according to a memorandum released by the Carnegie Commission.

The commission, an independent panel set up in 1988 with backing from the Carnegie Corporation, also says that the National Science and Technology Council (NSTC), the cabinet-level body responsible for coordinating science and technology across the administration, should concentrate on a smaller number of policy issues of importance to the president.

Since the start of the first Clinton administration in 1993, OSTP has been led by Jack Gibbons, previously director of the Office of Technology Assessment. OSTP is supposed to help Gibbons to advise the president on science and technology issues. Through the NSTC, it is also intended to coordinate science and technology policy across the federal government.

But the office has recently been criticized as lacking influence in the White House. Last month, Congressman George Brown (California, Democrat) told Science and Government Report, a Washington newsletter: "I don't think he [Gibbons] is winning the bat-



Gibbons: 'needs a chief of staff'.

tles over there on behalf of R&D."

The Carnegie Commission was set up in order to consider science policy issues under the joint chairmanship of William Golden, treasurer of the American Association for the Advancement of Science, and Joshua Lederberg, the Nobel prize-winner and former president of Rockefeller University in New York.

The commission completed its work last autumn, and issued its memorandum on OSTP as a parting shot to guide the new administration. It argues that the OSTP director's effectiveness would be "greatly increased" if he interacted more intensively on a day-to-day basis with the president's senior staff. The appointment of a chief of staff would give the director more time to do that.

It also thinks that OSTP has too many priority programmes, and says that the powerful White House Office of Management and Budget (OMB) is reluctant to let OSTP set 'crosscuts', which calculate government spending in areas of research across several government agencies.

The memorandum admits that OSTP

staff think that the relationship between the two offices has greatly improved. But it adds that "the view from OMB is more circumspect". It suggests a joint memorandum of understanding on the working relationship between the two offices.

Finally, the commission recommends that the NSTC's 60 or so working groups and sub-committees be pared back, and its resources concentrated on a small number of priority policy issues of concern to the president.

But it says that the overall structure for dealing with science and technology issues at the White House "is sound", and that the changes it suggests "are relatively minor".

Gibbons was in China last week with Vice-President Al Gore, and unavailable for comment. But Tim Newall, a spokesman for OSTP, said: "Overall we're pleased with the commission's endorsement of the steps we've taken to strengthen the science and technology policy making structure." The recommendations for change are "constructive suggestions which we'll take seriously".

But other observers were less charitable. The effectiveness of OSTP "has nothing to do with the structure, and everything to do with the person you have in there [as director] and their access to the president. All the rest is window dressing," says one physicist with extensive policy experience.

C.M.

MPs warn of threat to the science base

[LONDON] In its parting shot before the dissolution of the UK Parliament, the House of Commons Select Committee on Science and Technology warned that the pursuit by government departments of their individual agendas was threatening to undermine the strength of Britain's science base.

Two weeks ago, a similar warning was made in the context of the decision by the Ministry of Agriculture, Fisheries and Food (MAFF) to terminate funding for research at the Roslin Institute in Scotland that has led to the first successful cloning of an adult sheep (see *Nature* 386, 316; 1997).

The new warning comes in the committee's report on the activities of the Natural Environment Research Council, which was published last week. The report is based on a series of hearings over the past six months which concentrated in particular on NERC's involvement in research into climate change.

The committee was reluctant to endorse the criticism of several academic witnesses who complained at these hearings that NERC had shifted its research agenda too enthusiastically towards goals identified in the recent Technology Foresight exercise, to the detriment of research not lying within identified priority areas.

The committee agreed that the research community has "some grounds for its concern about NERC's commitment to basic research". But it accepted the research coun-

cil's reassurance that the decline in responsive research is likely to be reversed, adding that, "from the evidence available to us at this stage, NERC's approach seems entirely reasonable".

The committee is less generous towards the government, however, claiming that individual departments are often led to take actions whose consequences — often unintended — shift an additional financial responsibility on to the research councils, and so effectively reduce the amount of science funding available for other activities.

One example cited by the committee is a decision by the Meteorological Office to reduce its funding for flights of a C-130 aircraft that had been converted to make atmospheric measurements, following the withdrawal of financial support from the Ministry of Defence.

John Krebs, the chief executive of NERC, admitted to the committee that he feared that new financial arrangements introduced by the Met Office to recover the costs of the C-130 flights made it possible "that the UK core capacity in atmospheric science will decline in the next five years". That is a fear that is widely shared among atmospheric scientists.

Another potential threat to NERC's budget highlighted by Krebs was the possibility that the research council might be asked to contribute directly to the support of the

European Space Agency's Earth Observation programme. At present, this support is being provided directly by the Department of Trade and Industry through the British National Space Centre.

"It is easy for each department to pursue the course which will increase its ability to concentrate upon functions it regards as its core, and maximize the resources it can devote to that core," says the select committee in its report. But it adds: "The government as a whole must be aware of the danger that even though the headline figure of the science budget may remain stable, a wider range of responsibilities will mean that it is spread more thinly and less effectively."

Parallel conclusions had been reached by a separate report on the research council system, published two weeks ago, which cited some harsh criticism heard in evidence from Ray Baker, secretary of the Biotechnology and Biological Sciences Research Council, of the poor level of scientific understanding of some government officials formally responsible for distributing research funds at MAFF.

The committee ended this report by urging its successor committee, which will be appointed after the general election on 1 May, to carry out an examination of the relationship between the science budget and the expenditure by individual departments on science and technology. □

US National Science Board seeks a wider role in policy-making

[WASHINGTON] The US National Science Board (NSB) has endorsed an audacious plan to broaden its activities beyond the oversight of the National Science Foundation (NSF) and to offer broad science policy advice to both the administration and the Congress.

The board last week endorsed a plan to "expand its attention to larger national science and engineering policy issues". This would involve developing recommendations for how the government should set science priorities, and providing more analysis of the effectiveness of science policy in its biennial report, *Science and Engineering Indicators*.

The change has been instigated by Richard Zare, professor of chemistry at Stanford University and the new chairman of the science board (see *Nature* 382, 286; 1996). Zare promised that the board would address a "dysfunction" which, he said, was preventing the government from receiving clear advice from scientists on priorities. "The science policy community seems to confront the question of setting priorities again and again without resolving it," he said.

The NSB was set up at the same time as the NSF, shortly after the Second World War, primarily to monitor the activities of the science agency, which funds most non-biomedical university research in the United States. Some board members expressed misgivings about how deeply it should involve itself in the work of government agencies beyond its current role.

M. R. C. Greenwood, for example, chancellor of the University of California at Davis and a former senior official in the White House Office of Science and Technology Policy (OSTP), while voting for the plan, warned the board against talking about programmes in other agencies. She said that it "would stand a chance of losing its credibility" if it tried to participate in the government's annual budget process.

But Zare believes that it will be possible for the board to exercise wider influence — as it is legally entitled to do, under its 1950 founding statute — by expanding on the existing data collection work that it does in preparing *Science and Engineering Indicators*. But he is also aware of the

potential pitfalls. Last August, after chairing his first meeting of the board, Zare acknowledged that pursuing his new agenda would be "politically difficult and may cost the NSF support in the Congress".

Last week's meeting of the NSB also endorsed new procedures for the merit review of NSF grant proposals, along the lines suggested by a task force last year (see *Nature* 384, 399; 1996). The reforms will take effect in October, and will require referees to assess all proposals on the basis of their 'broader impacts' on, for example, education and society, as well as their intellectual merit.

In addition, the board announced the establishment of two major new supercomputing centres and the closure of two old ones. The new centres, which will operate as distributed facilities involving many partner institutions, will be based around the University of Illinois at Urbana-Champaign and at the University of California at San Diego.

The old centres, at Pittsburgh and at Cornell University, New York, will be phased out over the next two years. **Colin Macilwain**

Hubble's new camera hit by problem with solid nitrogen coolant

[WASHINGTON] Engineers checking the refurbished Hubble Space Telescope have found a problem with the new Near Infrared Camera and Multi-Object Spectrometer (NICMOS) that may reduce its useful lifetime. The glitch has temporarily left the wide-field camera on NICMOS, one of three on the instrument, unable to focus properly.

The probable culprit is solid nitrogen used to cool the instrument's infrared detectors — the cryogen has expanded to the point where it has pushed a cold part of the Dewar flask in which it is stored into contact with a warmer part. Although the nitrogen is boiling away at an excessive rate, engineers hope that will ease the thermal contact and bring the camera's focus back to normal.

NICMOS was due to last four-and-a-half years, with its lifetime limited by the supply of cryogen. Project officials will not know for several weeks how much of the nitrogen will be lost, but hope to make up for any damage by rearranging observing schedules.

Indian space scientists 'seduced by commerce'

[NEW DELHI] The Indian government has been urged by the parliamentary committee on science and technology to take immediate steps to halt the departure of scientists and engineers from the country's space programme for more lucrative jobs in the private sector, particularly the computer industry.

The committee's comments follow a review of the activities of the Indian Space Research Organization, which concluded that the space programme had slowed down due to "young scientists and engineers leaving en bloc". As many as 560 scientists and engineers have quit the organization in the past six years, with 135 leaving in 1995–96 alone.

CJD experts warn of transfusion dangers

[PARIS] Blood from people at risk from transmissible spongiform encephalopathies should not be used for transfusion, according to those attending an international meeting of over 50 experts on Creutzfeldt-Jakob disease (CJD) convened last week by the World Health Organization. Three risk groups of donors were identified: people having a transmissible spongiform encephalopathy; those who received human growth hormone during the 1980s when the product was extracted from cadavers or dura mater transplants; and members of families where a case of CJD had occurred. Paul

Brown, a researcher at the US National Institutes of Health, presented results to the meeting suggesting that transmission could occur via blood in mice.

Czech academy reports successful slimming

[PRAGUE] The president of the Czech Academy of Sciences, Rudolf Zahradník, last week reported to Vaclav Havel, the president of the Czech Republic, on the completion of the academy's four-year post-Communist reorganization. Since 1993, the academy has been considerably reduced in size, with the closure of 21 institutes and a halving of staff numbers to 6,400.

Slimmer is better, according to Zahradník. "In terms of publications, the academy's research is more efficient than before", he says. As the final step in restructuring, the academy's general assembly last week voted to separate its two functions. A new independent Czech Learned Society is to be created, leaving the academy the sole task of running the 59 basic research institutes.

Marconi collection is withdrawn from sale

[LONDON] Following a series of protests, the collection of papers and radio equipment originally belonging to Guglielmo Marconi is to be withdrawn from sale and given to the Science Museum in London (see *Nature* 385, 669; 1997).

The collection, valued at £3 million (US\$4.8 million), will be made available for public access in Chelmsford, Essex, where Marconi built the world's first radio factory a century ago. This decision results from a deal struck between the electronics firm GEC-Marconi, the Science Museum, Essex County Council and Chelmsford Borough Council.

GEC-Marconi had planned to auction the 250 items of equipment and 750 letters and documents. Those opposed to the sale included Marconi's daughter Elettra.

Switzerland rejects ban on gene-altered food

[BERN] Switzerland has opened its doors to the immediate import of food containing gene-altered ingredients. It has rejected an appeal to ban the import of gene-altered soya developed by the US company Monsanto, despite protests from consumers and environmental groups. In line with draft directives of the European Union, of which it is not a member, Switzerland will now allow sale of such foodstuffs provided they are clearly labelled as containing gene-manipulated products.

The decision by the Swiss Federal Department of Domestic Affairs on Monsanto's soya means that Swiss chocolate

manufacturers may now use gene-manipulated lecithin in their products. But the decision came just too late for some manufacturers which were caught — ironically just before Easter — using the product when it was still banned. They had to recall 500 tonnes from the market. Toblerone, which had to recall 350 tonnes, says it will no longer use gene-manipulated lecithin for chocolate for the Swiss market, because consumers clearly do not want it.

US academy to study fusion reactor plans

[WASHINGTON] The US Department of Energy has asked the National Academy of Sciences to study the scientific value of the proposed International Thermonuclear Experimental Reactor project and to report back by December 1997. DoE wants an assessment of whether the proposed magnetic fusion machine will attain its technical goal of sustained ignition, and of how much new scientific knowledge will be gleaned from the project.

The outcome of the academy study is expected to influence strongly the level of support the United States will offer for the project, if it goes ahead after the design phase is completed in mid-1998. A separate technical assessment of the proposed reactor design being undertaken by the DoE's Fusion Energy Sciences Advisory Committee will be published later this month.



The comet Hale-Bopp, as it has been seen by millions around sunset in the Northern Hemisphere in recent weeks, when it has been easily visible to the naked eye. This photograph was taken by Alan Fitzsimmons of the Department of Pure and Applied Physics at Queen's University in Northern Ireland. It was taken near the Giants Ring, south of Belfast, and used a 15-second exposure with a 50-mm lens.

Advantage of knowing nature's secrets

Sir — The epoch-making results of Wilmut *et al.*¹ describing the cloning of adult sheep once more illustrates the perpetual problem of equating scientific endeavour with the public interest. I noted with interest the media frenzy which greeted these findings together with the subsequent public discussion of the likelihood and implications, both practical and moral, of future cloning of adult humans. We scientists are intrigued and driven by every new discovery no matter how small. We also, however, have an obligation to relay this information to the public in a manner that creates a suitable environment for informed discussion to take place.

In this respect, I was reminded of the cautionary words of Pierre Curie on the occasion of his address to the Swedish Academy of Sciences in 1903 when he and Marie Curie received the Nobel Prize for Physics: "...one can ask if humanity is at an advantage in knowing nature's secrets, if it is mature enough to make use of them or if this knowledge might not be harmful to it"². I, for one, agree with *Nature's* decision to publish the results of Wilmut *et al.*, and concur with Curie's conclusion that "...humanity will derive more good than bad from new discoveries"².

Michael J. Taggart

Physiological Laboratory,
University of Liverpool,
Liverpool L69 3BX, UK

Sir — Your recent News article³ on the reaction to sheep cloning suggests that "scientists agree that cloning of humans would be unethical". Indeed, a certain Jeremy Rifkin goes so far as to suggest that penalties for human cloning should be on a par with those for "rape, child abuse and murder".

There are two issues that Rifkin and his band of crusaders should bear in mind. First, history shows that scientific knowledge is used and developed, no matter what ethical considerations are involved. For example, despite the many cries to "Ban the Bomb", nuclear weapons have been developed by all the major (and some minor) powers. Second, one could imagine a situation in which human cloning might be the only option available to recover a human life. Should Rifkin deny the option of cloning to the parents of a young child fatally injured in an accident, if viable cells could be recovered from the body?

I am not aware of any conclusive poll that confirms that "scientists agree that cloning of humans would be unethical". Moreover these issues are too important to

be left to the champions of the knee-jerk reaction highlighted in your news article.

Mike Fainzilber

Molecular Neurobiology Laboratory,
Karolinska Institute BZ-Lab.,
Doktorsringen 12A,
S17177 Stockholm, Sweden
e-mail: michael@cajal.mbb.ki.se

Sir — The publication of the experimental protocol through which Dolly came to life¹ created a considerable stir in the media. In this context, the arguments put forward in your leading article⁴ in favour of a mere moratorium rather than total and definitive banning of this type of experiment in humans are not acceptable. More precisely, it is unacceptable to let people think that in certain circumstances human beings could be brought to life by means of somatic cell cloning, because the essence of humanity lies in the uniqueness of each of its members, resulting from the unique recombination of two equally unique genomes.

In consequence, the only solution is to pronounce a universal ban on human cloning. When this is done — and only then — can one envisage a one-year moratorium to define the conditions in which cloning experiments involving human cells could be performed, but for the sole advance of knowledge and the ultimate prospect of growing *in vitro* differentiated cells or organs intended for transplant therapy.

Francis Galibert

University of Rennes 1,
UPR 41,
2 avenue du professeur Leon Bernard,
35043 Rennes Cedex, France

Jacqueline Godet

Centre de Génétique
Moléculaire et Cellulaire,
CNRS/University of Lyon,
69622 Villeurbanne, France

Jean-Claude Kader

Laboratoire de Physiologie
Cellulaire et Moléculaire,
URA CNRS 2135, Université Paris 6,
75252 Paris Cedex 05, France

Jean-Antoine Lepasant

Institut Jacques Monod,
CNRS/Université Paris 7,
75251 Paris Cedex 05, France

Sir — In the Human Fertilisation and Embryology Act 1990 and the Animal (Scientific Procedures) Act 1986, the United Kingdom has a regulatory framework which was already in place and which was applied to the work on the cloning of sheep at the Roslin Institute.

Nature has reported accurately and

systematically the work of the House of Commons Select Committee on Science and Technology, and in particular our report on human genetics, in which we recommended the setting up of the Human Genetics Advisory Commission as a standing body to report publicly on all aspects of human genetics and its applications, including ethical, legal and social aspects. The commission was meeting for the first time on the very day *Nature* published the Letter reporting the results of the successful cloning of a sheep¹.

It was therefore disappointing that your leading article⁴ concluded by saying it "is shaming" for "politicians only now to be requesting guidance about what appears in today's *Nature*".

We are however confident that our further current inquiry into cloning, and the work of our successor committee in the next parliament, will continue to be well reported in *Nature*. It is an important channel of communication not only within science, but also with governments and politicians about science and its impact on society.

Giles Shaw

(Chairman)

Jeremy Bray

Select Committee on Science and Technology,
House of Commons,
London SW1A 0AA, UK

● The leading article actually refers to "a US president and other politicians..."
— Editor, *Nature*.

Sir — I write to comment on the picture of Dolly, the cloned sheep, on the cover of the 27 February issue. Superficial inspection of this photograph reveals that Dolly unexpectedly has one black hoof. Could it be that the cytoplasm of the enucleated Scottish Blackface egg contains unknown genetic information that specifies the left rear hoof? Or, is there some other Photoshop-like explanation for this maternally derived appendage (see Figure 2, page 812)?

The possibilities are especially interesting coming on the heels (hooves?) of the recent report of a fowl-up when quail neural cells are transferred into chicken brains.

Sidney Strickland

University at Stony Brook,
Stony Brook, New York 11794-8651, USA.
e-mail: strickland@pharm.sunysb.edu

1. Wilmut, I. *et al.* *Nature* **385**, 810–813 (1997).

2. Quinn, S. *Marie Curie. A story* (Mandarin, UK, 1995).

3. Masood, E. *Nature* **385**, 757 (1997).

4. *Nature* **385**, 753 (1997).

Unfair treatment

Sir — I completely agree with Franklin D. Rumjanek about acceptance of papers from developing countries by leading Western scientific journals (*Nature* 384, 509; 1996). I have had several bad experiences and I feel that many of my papers have been arbitrarily rejected.

There is one more misdeed indulged by many Western scientists. Published papers (even in journals published in the West) are never cited when the editors of the relevant journals are approached — they send a bland reply saying that the journals policy is not to correct faulty citations.

The same fate has befallen several of my colleagues. Most of us remain silent because we know we will not get any justice.

V. D. R. Nathan

40 Muthiyalu Chetty Street,
Vepery, Chennai 7,
600007, India

Sir — Rumjanek presents the difficulties of publishing a manuscript produced by his group in Brazil, although the quality of the work was essentially the same as papers he published while he was a postdoc in England. He suggested that the referees could be influenced by the geographical

location of the laboratory responsible for the submitted manuscript. As I am working as a postdoc in France and thinking of going back to Brazil in the near future, I am conscious of this problem. I should like to suggest a protocol to the editors of scientific journals that could eliminate or diminish some of the problems encountered by Rumjanek and other scientists.

As the authors of a manuscript don't know the referees, why don't the referees who evaluate a given manuscript do it without knowing its origin and authors?

Hernan Terenzi

Unité d'Expression
des Genes Eucaryotes,
Institut Pasteur,
28 rue du Dr Roux,
75724 Paris, France
e-mail: hterenzi@pasteur.fr

Sir — I have a small contribution to the recent correspondence on national prejudice in reviewing. I spent my first 28 years in England, and was educated at an old-fashioned high school where I was always near the top of the class in English language and literature. Now I live and work in Germany. Last year, a paper I had submitted was returned by the journal concerned complaining that it was "strewn with spelling mistakes and grammatical errors". Neither accusation was true, according to both my

(admittedly British) spelling checker and a professional linguist friend. Nevertheless, the editor advised me to seek the help of "a colleague whose mother tongue is English".

Christine Clayton

ZMBH,
Im Neuenheimer Feld 282,
D69120 Heidelberg, Germany
e-mail: cclayton@sun0.urz.uni-heidelberg.de

Life imitating art?

Sir — I was seized with a sense of *déjà vu* while reading of the summary dismissal of Dr Mark Hughes from his position as head of preimplantation and prenatal diagnosis at the US National Institutes of Health (NIH) and the disavowal by the NIH and Georgetown University of any knowledge of his actions (*Nature* 385, 190; 1997).

I was transported into Rick's Café Americain in the film classic *Casablanca*. Ordered by Major Strasser of the Third Reich to close the establishment, Louis blows his whistle and announces that Rick's is closed because "I am shocked to learn that gambling is going on here". Precisely at that moment, he is handed his night's roulette winnings.

Hughes was recruited from his highly successful laboratory at Baylor University to head the NIH's programme of

Obtaining **reliable**
PCR results means
using **Ready-To-Go**
PCR Beads

Terri Davis is a cellular
biochemistry technician
working in New York, NY.



preimplantation genetics. Unfortunately, a rider attached to the compromise federal spending bill at the insistence of the Republican majority forbade all federal funding of research involving human embryos. The ultimate irony is that preimplantation genetics was developed to provide an alternative to aborting genetically abnormal fetuses, a goal consistent with the Republicans' opposition to abortion. Yet this law prohibited federally funded research in this field. Faced with this repressive law, Hughes endeavoured to move his preimplantation activities off the NIH campus and into the private sector to comply with the law while continuing his ground-breaking research.

After allegations by a postdoc, it became public knowledge that Hughes was doing preimplantation genetics at his off-campus site. Instead of support from the scientific community and outrage at government intrusion into genetic research, NIH and Georgetown scientists and administrators fell over themselves to express "shock" that preimplantation diagnosis was going on. Shocked that a brilliant scientist recruited to develop the field of preimplantation genetics was caught red-handed actually doing preimplantation genetics! Why not support the man they recruited? Instead of feigning ignorance and suggesting unethical behaviour, why not tell the public

that Hughes made extraordinary efforts with restrictive federal laws while still offering couples a chance to avoid aborting abnormal fetuses?

In *Casablanca*, Rick and Louis eventually team up to defy the oppressors. Unfortunately that kind of courage was not evidenced in recent events. Hughes was abandoned and hung out to dry.

Hughes' dismissal and abandonment is a profound and disturbing example of the continued politicization of American science. Soon only politically correct projects will be funded and carried out, and science will become just a pawn in the partisan political agenda.

Charles M. Strom

Illinois Masonic Medical Center,
836 Wellington Avenue,
Chicago, Illinois 60657, USA

Not too old at 30

Sir — Why are advertisements for PhD students specifically directed at those who are "under 30" or a "young student"? I am surprised that a journal such as *Nature* should support such discrimination. Indeed, it may conflict with equal opportunities legislation.

I am a mature student, one of many in universities. We mature 30–40-year-old budding scientists who are dedicated to our

subject would love an opportunity to undertake PhD research. But not only are we constantly reminded of our lack of 'neuronal plasticity' by the industrial sector, but it seems that advertising for basic research is allowed blatantly to reinforce this view. Mature students with good degrees and prizes get the impression that further research must be through low-profile, ill-funded and very low paid studentships. This ensures that after finishing a PhD we are specialized in a field with little applicability to industry. As for academic institutions, expanding industrial links with universities also prohibit employment of those specialized in low-profile areas.

It is more than disheartening for mature students, often with a flair for science, to be discriminated against in this way. Heaven forbid if we are female too! Perhaps *Nature* should set an example by adopting a policy of ensuring that those placing advertisements realize the concept of equal opportunities.

Gina Clayton

c/o Biological Laboratories,
University of Kent,
Canterbury, Kent

● Stimulated by this letter, *Nature* intends in future to discourage advertisers from discriminating on the grounds of age.

It is already illegal in some countries.

— Editor, *Nature*.

There are many opportunities for introducing contaminants into your PCR* reaction which could jeopardize the reliability of your results. Preparing individual PCR reactions every time you want to perform PCR is not only time consuming, but allows for pipetting errors. Sharing a master PCR mix with others could also compromise the integrity of the PCR reaction, resulting in reduced reliability. Pharmacia Biotech has applied the Ready-To-Go® bead technology representing pre-mixed, pre-formulated, ambient-stable, single-dose reactions to typical PCR reactions, yielding Ready-To-Go PCR Beads.

New ambient-stable Ready-To-Go PCR beads: just add primers, DNA and water

Obtaining reliable PCR results has never been so simple and reliable. You begin by removing a 0.5 ml tube containing an optimized PCR bead from the foil pouch. Next, watch the bead dissolve in seconds as you add DNA, primers, and water. Then, place the tube into a thermocycler and turn it on. That's all it takes to obtain reliable PCR results. There's nothing like it on the market today!

The pre-formulated PCR beads deliver unprecedented reliability and reproducibility for amplifying DNA fragments from a variety of DNA templates, such as plasmid, cDNA, genomic sources and viral DNA. This high level of performance is achieved because every time you perform PCR, a fresh reaction bead is used. Fewer steps means minimizing the risk of contamination, and increasing the reliability of your results. Each batch of Ready-To-Go PCR Beads is function-tested, ensuring lot-to-lot consistency.

To find out more, give us a call: 1 (800) 526-3593 in the USA; +81 3492 6949 in Japan; +46 18 16 50 11 in Europe and the rest of the world. Or visit us on the Internet: <http://www.biotech.pharmacia.se>.

**Pharmacia
Biotech**
Uppsala, Sweden. (And the rest of the world)

READER ENQUIRY NO. 606

A Ready-To-Go PCR Bead: each is optimized for PCR and contains buffer, nucleotides, and Taq DNA Polymerase. (Magnified 270%).

*PCR is a patented process of Hoffman-La Roche



Scientists are still keeping the faith

Although the suggestion eighty years ago that four in ten scientists did not believe in God or an afterlife was astounding to contemporaries, the fact that so many scientists believe in God today is equally surprising.

Edward J. Larson and Larry Witham

To measure the strength of religious belief in an era of ascendant science, the eminent researcher James Leuba conducted a landmark survey in 1916. He found that 60 per cent of 1,000 randomly selected scientists did not believe in a God, and predicted that such disbelief would increase as education spread¹. To test that prediction, we replicated Leuba's survey as exactly as possible. The result: about 40 per cent of scientists still believe in a personal God and an afterlife. In both surveys, roughly 45 per cent disbelieved and 15 per cent were doubters (agnostic).

Eighty years ago, Leuba wrote that scientific knowledge would demand "a revision of public opinion regarding the prevalence and future of the two cardinal beliefs of official Christianity". He asserted: "The essential problem facing organized Christianity is constituted by the widespread rejection of its two fundamental dogmas¹." Though a noted psychologist, Leuba misjudged either the human mind or the ability of science to satisfy all human needs. Such is the risk of making historical predictions.

In the intervening years, religious belief has become more diverse. But, to the extent that both surveys are accurate readings, traditional Western theism has not lost its place among US scientists, despite their intellectual preoccupation with material reality. But it should not be forgotten that Leuba's finding of widespread disbelief among US scientists was astounding in its day. And today, even more than in 1916, most scientists have no use for God or an afterlife.

Speaking for this majority, one 1996 respondent to the question about desiring immortality said, "It is pointless to desire the ridiculous". Leuba's interest in psychology prompted him to insert this secondary question, asking whether respondents who did

not believe in personal immortality nevertheless desired it. More than any other of his questions, it addresses the tension between intellect and emotion in some scientists. In 1916, 73 per cent of nonbelievers nevertheless desired immortality¹. That hope has dwindled over 80 years (see Table 1).

Despite the stability in the overall proportion of believers and disbelievers, there has been a significant shift in views held by the three professions surveyed — mathematics, biology and physics/astronomy. The 1996 survey showed that mathematicians are most inclined to believe in God (44.6 per cent). And although biologists showed the highest rates of disbelief or doubt in Leuba's day (69.5 per cent), that ranking is now given to physicists and astronomers (77.9 per cent).

Higher belief among physicists in Leuba's survey might have been expected at a time when leading physicists such as Lord Kelvin, Robert Millikan and Sir Arthur Eddington publicly defended religious belief². We also expected a higher proportion of belief among physicists and disbelief among biologists, particularly with some prominent astrophysicists in the 1990s entertaining the anthropic principle and certainly the advent of Big Bang cosmology — whereas most prominent biologists still stick hard by darwinian naturalism³. But we were wrong.

Leuba's documentation of disbelief had a political impact, particularly his data on the decline in belief among US college students. The populist Democratic politician Williams Jennings Bryan and some conservative Christians seized on Leuba's data in the 1920s to show the social evils of modernism, and its ultimate impact on the country's morality. They accused academic scientists of leading college students into disbelief⁴⁻⁶.

Although Leuba's data caused a scandal, the same findings for 1996 may do just the opposite. Today, many people presume that

scientists are far less likely to believe in the supernatural than the general population, so religious Americans will doubtless be pleased to know that as many as 40 per cent of scientists agree with them about God and an afterlife. Nowadays, the US scientific enterprise is fixated on the battle for federal funding, spawning calls for the "civic scientist" or "scientist as citizen", ultimately to promote the cause of science to taxpayers. If more than one-third of scientists hold beliefs dear to many conservative Americans, the knowledge that scientists are "just like us" may evoke in them a sympathetic response.

Compared to the technology used in modern surveys, Leuba's effort was quaint. Nevertheless, his was among the earliest efforts to apply the science of statistics to sociology. In two separate mailings, he sent his survey to a total of 1,000 scientists drawn randomly from the 1910 edition of *American Men of Science*. He received about a 70 per cent response. We randomly drew 1,000 names from the current edition of *American Men and Women of Science*; our response was about 60 per cent. We stuck to Leuba's apportionment: half biologists and a quarter each in maths and physics/astronomy.

Oddly, in Leuba's survey about 20 per cent of the scientists who did not believe in God nevertheless reported a belief in personal immortality. We obtained a more logical response in that belief in God and in immortality were almost always held together. It is unclear from Leuba's research just what scientists earlier in the century thought of as life beyond death, even with no God — though it might have meant the conservation of energy or, in Cicero's terms, being remembered by later generations. The questions go to the issue of theological definitions. Leuba recognized this, but kept the format of his questionnaire to a simple two questions of three choices each, saying that "in attempting to refine, I should probably have made matters worse".

Leuba viewed the great challenge to science as orthodox Christianity, which is why he defined God in very conventional terms: "A God to whom one may pray in expectation of receiving an answer". We believe that because such traditional tenets still prevail in American culture, retaining Leuba's 1916 definition of God — hearing prayers and giving immortality — still gives the best simple question.

If respondents in Leuba's time did not agree with his survey in general — one respondent said, "This is a lot of damned rot!" — we received unsolicited comments that the definition of God did not allow for enough variation. "Why such a narrow

Table 1 Comparison of answers to questions in 1916 and 1996 surveys

Topic of question	1916	1996
Belief in personal god		
1. Personal belief	41.8	39.3
2. Personal disbelief	41.5	45.3
3. Doubt or agnosticism	16.7	14.5
Belief in human immortality		
1. Personal belief	50.6	38.0
2. Personal disbelief	about 20*	46.9
3. Doubt or agnosticism	about 30*	15.0
Desire for immortality		
1. Intense	34	9.9
2. Moderate	39	25.9
3. Not at all	27	64.2

*Results published graphically by Leuba. All figures are percentages.

A STATISTICAL ENQUIRY

Conflicting statements are confidently made regarding (whether scientists hold a) belief in God and in personal immortality. Nevertheless, sufficient data are not extant to support any such supposition.

The accompanying questions are sent to 1000 persons taken by chance from those listed in "American Men (and Women) of Science", in the hope of securing statistics valid for this group. The condition of success is that all those addressed respond. No satisfactorily definite conclusions could be drawn if many of those addressed refused or neglected to answer.

It will take you only a few seconds to make a mark by every statement true for you. Please do it, if at all possible, on receipt of this paper and return it in the enclosed stamped envelope.

A. CONCERNING THE BELIEF IN GOD.

-1. I believe in a God in intellectual and affective communication with humankind, i.e. a God to whom one may pray in expectation of receiving an answer. By "answer" I mean more than the subjective, psychological effect of prayer.
-2. I do not believe in a God as defined above.
-3. I have no definite belief regarding this question.

B. CONCERNING THE BELIEF IN PERSONAL IMMORTALITY,

i.e. the belief in continuation of the person after death in another world.

1. I believe in:
 - a. personal immortality for all people
 - b. conditional immortality, i.e. for those who have reached a certain state of development.
2. I believe neither in conditional or unconditional immortality of the person.
3. I have no definite belief regarding this question.
4. Although I cannot believe in personal immortality, I desire it:
 - a. intensely
 - b. moderately
 - c. not at all

Questions of belief: how Leuba phrased his 1916 survey of American scientists.

definition [of God]?" asked one of our respondents, writing in the survey margin. "I believe in God, but I don't believe that one can expect an answer to prayer."

Surveys in the past decade have shown that religious belief, professed by 93 per cent of Americans, has become more diverse. When Americans are asked to define 'God', a quarter opt for something other than a conventional theistic deity. They see 'God' as higher consciousness (11 per cent), full realization of personal potential (8 per cent), many gods (3 per cent) or everyone as their own god (3 per cent). A more robust sampling by Leuba and by us would doubtless have provided more confidence that the results reflect what scientists really believe. Yet our findings do corroborate a large survey done in 1969 by the Carnegie Commission, asking 60,000 professors in the United States questions such as "how religious do you consider yourself?". The commission found that 34 per cent of physical scientists were "religiously conservative" and about 43 per cent of all physical and life scientists attended church two or three times a month — on a par with the general population⁸.

Due to the dramatic expansion in the US scientific community since 1916, our survey polled only about 3 per cent of the biological and physical scientists and mathematicians listed in the 1995 *American Men and Women of Science*⁹. From his listing, Leuba reached more than 20 per cent of these groups¹⁰. In Leuba's day, the editors of the reference book went through the painstaking process of deciding who were "great scientists" compared to ordinary, and an asterisk was put by those names¹⁰. Thus, Leuba could analyse belief among the great compared to "ordinary". He found a higher amount of disbelief among great scientists, but this was a category we could not test — the asterisks do not appear in modern editions.

Another way to make this distinction would be to compare 'research' scientists to teachers or those working in industrial technology, for example. Names listed in *American Men and Women of Science* are judged by awards, citation or recommendation. The editors tried for several years to highlight 'research' scientists, but abandoned the effort for lack of consistent responses from the tens of thousands of people listed. Our

best indication, therefore, comes from research by the National Science Foundation, which estimates that about 11 per cent of scientists are involved exclusively in basic research¹¹. Whatever the stature of respondents, the positive unsolicited responses to our survey were more plentiful than Leuba's. All of our respondents at least attempted to answer the questions, whereas Leuba had a nearly 10 per cent rejection rate, with comments such as "How is it possible for a sane student to answer these questions?"

Under conditions of anonymity, we offered to send the results to the subjects of our survey, and their responses gave us an inkling of the professional landscape with which we had made contact. Letters came from well-known private universities such as Chicago and Johns Hopkins, but more so from land-grant schools. We heard from scientists at national research centres, as well as from technical institutes, medical colleges and commercial labs — and from one Nobel laureate.

The persistent interest in, even struggle with, religious questions among scientists was poignantly captured in one handwritten letter from a Harvard professor. "When backed into a corner, as it were, by questions such as those on the survey, I have to come down on the side of non-belief", the scientist wrote. "This result for me, however (and possibly for others), is an unduly harsh picture. I try frequently to open my mind to an influence of what is good, and the 'subjective and psychological' effects of this can be quite profound, such that I am happy to make contact with the religious tradition by saying that I am praying to God."

Similarly, after indicating no desire for immortality, one of our respondents added wryly, "But it would be nice". □

Edward J. Larson is in the Department of History, University of Georgia, Athens, Georgia 30602, USA, and at the Discovery Institute, 1201 Third Avenue, Seattle, Washington 98101, USA (e-mail: edlarson@uga.cc.uga.edu). Larry Witham is at 3816 Lansdale Court, Burtonsville, Maryland 20866, USA.

1. Leuba, J. H. *The Belief in God and Immortality: A Psychological, Anthropological and Statistical Study* (Sherman, French & Co, Boston, 1916).
2. Brooke, J. H. *Science and Religion: Some Historical Perspectives* (Cambridge Univ. Press, Cambridge, 1991).
3. Templeton, J. M. (ed.) *Evidence of Purpose: Scientists Discover the Creator* (Continuum, New York, 1994).
4. Marsden, G. M. *The Soul of the American University: From Protestant Establishment of Established Nonbelief* (Oxford Univ. Press, New York, 1994).
5. Numbers, R. L. *The Creationists: The Evolution of Scientific Creationism* (Knopf, New York, 1992).
6. Larson, E. J. *Trial and Error: The American Controversy Over Creation and Evolution* (Oxford Univ. Press, New York, 1989).
7. Barna, G. *The Index of Leading Spiritual Indicators* (Word, Dallas, 1996).
8. Stark, R. & Iannaccone, L. *American Economic Review: Papers and Proceedings*, 433–437 (1996).
9. *American Men and Women of Science* 19th edn (Bowker, New Providence, New Jersey, 1995).
10. Cattell, J. M. (ed.) *American Men of Science: A Biographical Directory* 2nd edn (Science Press, New York, 1910).
11. National Science Board *Science and Engineering Indicators* (Washington: National Science Foundation, 1996).

Cutting complexity down to size

David I. Stuart and E. Yvonne Jones

The proteasome is the most complex eukaryotic macromolecular assembly yet seen in fine detail. The structure reveals completely unexpected mechanisms by which the proteasome neatly chops up unwanted proteins for disposal or display to the immune system.

Controlled degradation of proteins inside the cell is central to many processes, ranging from cell-cycle control and differentiation to the cellular immune response. The targeted protein is marked for destruction by the covalent attachment of a small protein called ubiquitin and, following ATP-driven unfolding, it is degraded in the secluded inner chamber of a large protein complex known as the 26S proteasome. The core and proteolytic chamber of the 26S proteasome are formed by the catalytic 20S particle. This is flanked at either end by the so-called 19S 'caps', which seem to be responsible for recognition and unfolding of the ubiquitin-tagged target protein (Fig. 1).

The eukaryotic 20S catalytic engine is composed of 14 different, but related, protein types which assemble, with absolute precision, into a single structure of 28 protein chains — a leviathan of the molecular world. It can now be seen on page 463 of this issue, where Groll *et al.*¹ report a crystallographic *tour de force*: they have solved the structure of the *Saccharomyces cerevisiae* 20S proteasome to atomic resolution. Discounting symmetrical systems such as viral capsids, it is the most complex structure yet analysed, showing that complexity is not in itself a barrier to understanding at the atomic level. The structure also throws existing ideas about the ubiquitous degradation process into question, and allows the authors to put forward some radical suggestions for how this deceptively complex molecular conglomerate might work.

The structure of the eukaryotic proteasome is so revealing because it is far more complex than that of an archaeobacterial cousin², which comprises just two types of subunit, α and β . The α -subunits do not seem to be catalytic, but they can self-assemble into seven-membered rings. Conversely, the β -subunits possess proteolytic activity but they cannot, alone, assemble. Together, the two components arrange themselves into a stack of four rings — two outer α -rings guarding an inner pair of β -rings. The eukaryotic proteasome maintains the same essential features, but it has diversified, presumably reflecting enriched biological functions. The exact sevenfold symmetry of

the particle is lost, and only twofold symmetry remains, so that there are seven different α - and seven different β -subunits.

The basic 20S architecture was revealed by studies of the archaeobacterial proteasome which, coming hot on the heels of analysis of the chaperone protein GroEL, provided a second example of a cavity that could hold unfolded protein³. But whereas in GroEL the unfolded protein is nurtured, increasing the chance of correct folding, in the proteasome its vulnerability is taken advantage of, and it is cut to length in the inner chamber. There are three key questions. Where does the unfolded protein enter the inner chamber?

How is it cut? What is the ruler that measures the distance between the cuts? The structure of the eukaryotic proteasome, solved by Groll *et al.*¹, suggests some answers.

The structure of the archaeobacterial proteasome indicated that unfolded peptide might be able to slip into the proteolytic chamber along the axis of sevenfold symmetry, through small openings between the α -subunits. But the authors show that in the eukaryotic proteasome these doors are firmly locked shut — the only apparent way in or out is through some rather small side windows. So how do proteins enter? *In vivo*, the 26S eukaryotic proteasome works as an ATP-driven machine, comprising the 20S catalytic engine and the 19S caps. Because of the position of these 19S structures, and by analogy with the observed opening in the archaeobacterial proteasome, it is tempting to speculate that unfolded proteins are fed into the interior of the eukaryotic 20S core through the centre of the α -subunit ring. Presumably, this occurs by the 19S-mediated opening of a gate, which is closed in the isolated 20S structure.

The second question is, what is the cleavage mechanism? Hydrolases — a broad clan that cut and clip away at substrates with the

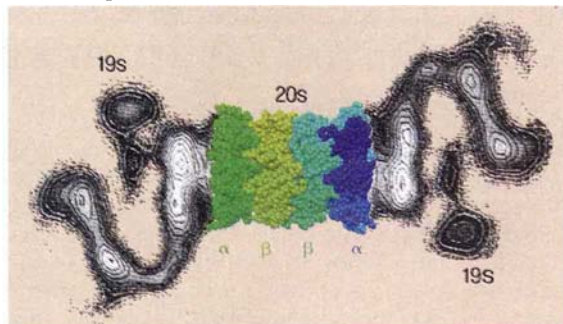


Figure 1 The 26S eukaryotic proteasome (M_r 2,000,000) is assembled from a 20S core, whose four-ring structure (two α - and two β -rings) has been determined by Groll *et al.*¹ using X-ray crystallography, and two 19S caps, shown here as an image based on the results of electron microscopy⁷.



Figure 2 The 28 subunits of the yeast proteasome are viewed with the quasi-sevenfold axis horizontal. One β -subunit has been moved upwards (grey shadows mark the original positions of its helices) to reveal the octameric amino-terminal extension of another (shown in red). The threonine residue that would cut at the carboxy terminus of this peptide in an active subunit is shown in grey, whilst the amino terminus (shown in blue) may have been cut by a different active site at the end of the adjacent helix. The conformation of the octameric peptide is strikingly similar to that characteristic of MHC-bound peptides.

aid of water — have been studied structurally for longer than any other group of enzymes⁴. Within this clan, the proteases were thought to employ a small number of relatively well-understood catalytic strategies. Then, in 1995, the report of the archaeobacterial 20S proteasome structure², along with an accompanying mutagenesis paper⁵, brought a new twist: they described a novel reaction mechanism that directly involves the amino-terminal threonine (Thr 1) of the β -chain. This mechanism is now known to be used by other hydrolases, which together form the 'N-terminal nucleophile' (Ntn) family⁶.

In the proteasome, the Ntn active sites line the internal cavity that is formed by the β -rings (Fig. 2). The structure of the eukaryotic proteasome, solved by Groll *et al.*¹, clarifies aspects of the Ntn reaction mechanism, in particular by revealing the presence of a water molecule that is positioned to act as a nucleophile. The reaction may proceed either through an acyl enzyme intermediate (involving attachment to the side-chain hydroxyl group of Thr 1) which is then deacylated by the water, or perhaps through the direct attack of the water molecule on the peptide bond to be cleaved.

This, however, is only the beginning of the story. All seven of the β -subunits are synthesized with extensions beyond the expected amino terminus. These extensions are involved in ensuring the correct assembly of the proteasome, but they must be cleaved to give functional active Ntn sites. Only three of the seven are cleaved. This pattern of cleavage makes sense if we assume that the cleavage is autocatalytic — those β -chains that either do not possess a catalytic Thr 1 or harbour a change in a nearby essential residue do not process properly, so they remain inactive.

But the picture is more complex than this. Whereas two types of inactive β -subunit show no cleavage at all, the amino-terminal extensions in two other types of β -subunit have been partially processed — but not by an Ntn site. Instead, the chain remains longer by eight or nine residues. These residual, amino-terminal tails are well ordered, being pinned to the walls of the internal cavity (Fig. 2). The position of the cleaved ends indicates that there is a second, previously unsuspected, active site at the carboxy-terminal end of an α -helix which is present in all seven types of β -subunit. So the complex machinery of the proteasome activates itself by a series of pre-processing events which, for an unknown reason, fail at different stages in most of the β -subunits, providing snapshots of how the normal *in vivo* cleavage of exogenous protein may take place using two types of active site.

This leads us to the question of the molecular ruler that measures the distance between the cuts inflicted on the proteins to be degraded. One of the functions of the proteasome is the generation of antigenic

peptide octamers or nonamers, which bind to class I molecules of the major histocompatibility complex (MHC). These peptide-MHC complexes are recognized by cytotoxic T cells during an immune response. The trimmed tails of the amino-terminal extensions would, if they were also cleaved at the Thr 1 site, be just the right length to form these peptide epitopes. So the two types of active site may define the two ends of the molecular ruler.

Specialized subunits replace two of the three active β -subunits in the standard proteasome on initiation of an immune response. The specificity of pockets adjacent to the Ntn sites, particularly in one of these specialized subunits, equip these sites to cleave selectively such that the generated peptides possess carboxy-terminal residues that are tailored to bind MHC class I. The second site would cleave to produce the amino terminus of the peptide. There is no pocket associated with this site, indicating that cleavage is less specific here. This would agree with the lack of bias in the type of amino-terminal residue that is required for MHC binding. Furthermore, the conformations of the partially processed 8–9-residue

tails in the structure of Groll *et al.*¹ bear a striking resemblance to the peptide epitopes that are presented in the binding groove of a class I MHC molecule. Could we be seeing not just a molecular ruler but also a device to select peptides which, when unfolded, favour the particular conformation that is required for binding to MHC class I?

The structure of the 20S proteasome core suggests plausible answers to the main functional questions and provides a detailed model for this enormous component (M_r 700,000). It also provides a potential key to unlock the complexities of the 26S particle (M_r 2,000,000), once X-ray or electron diffraction data can be obtained. □

David I. Stuart and E. Yvonne Jones are in the Laboratory of Molecular Biophysics and Oxford Centre for Molecular Sciences, Rex Richards Building, South Parks Road, Oxford OX1 3QU, UK.

1. Groll, M. *et al.* *Nature* **386**, 463–471 (1997).
2. Lowe, J. *et al.* *Science* **268**, 533–539 (1995).
3. Weissman, J. S., Sigler, P. B. & Horwich, A. L. *Science* **268**, 523–524 (1995).
4. Blake, C. C. F. *et al.* *Nature* **206**, 757–761 (1965).
5. Seemüller, E. *et al.* *Science* **268**, 579–582 (1995).
6. Brannigan, J. A. *et al.* *Nature* **378**, 416–419 (1995).
7. Yoshimura, T. *et al.* *J. Struct. Biol.* **111**, 200–201 (1993).

Supernovae

SN1987A enters its second decade

Richard McCray

On 23 February 1987, SN1987A came howling into the world of astrophysics, the brightest supernova since SN1604 was seen by Kepler. Astronomers were astonished that its parent was a blue giant star, startled by its triple-ring system, and disappointed not to find the pulsar that should have been born during the observed neutrino flash. Now, at age ten, SN1987A has settled into a relatively tranquil childhood, but we are beginning to see hints of a stormy adolescence. On the tenth anniversary of the discovery of SN1987A, astronomers met* in La Serena, Chile, to assess our understanding of this extraordinary event.

In some respects, observations of SN1987A have confirmed our theoretical picture of type-II supernovae, which are supposed to be driven by the stellar core collapsing into a neutron star. The neutrino flash confirmed the core-collapse; and the detection of γ -ray lines and infrared emission lines confirmed that the exponential decay of the supernova light was because the remnant was being heated by radioactivity from isotopes made in the explosion — 0.07 solar masses of ^{56}Co and 0.003 solar masses of ^{57}Co .

*Supernova 1987A: 10 Years After, La Serena, Chile, 23 February – 1 March 1997. Abstracts are at http://www.ctio.noao.edu/SN1987A_conf/index.html.

But many puzzles remain. Intensive efforts have yielded no clear evidence of a central neutron star or black hole. Independent observations have failed to confirm the observation of a 2.1-millisecond optical pulsar reported by J. Middleditch (Los Alamos Natl Lab.). To have escaped detection, the central compact object must have a total luminosity less than a few hundred times the Sun's, and far less than that of the 943-year-old pulsar in the Crab Nebula. The opacity of the infalling gas may be higher than previously thought, thus limiting the luminosity below detectability (C. Fryer & P. Pinto, Univ. Arizona). The compact object may elude detection for decades, especially if it happens to lie behind one of the dust clouds in the expanding debris.

While there is general agreement that the explosion must have been driven by neutrino heating, details of the mechanism are uncertain (A. Burrows, Univ. Arizona). The collapsing proto-neutron star is almost certainly convective at the neutrino sphere, where the neutrinos can stream freely outward, but it is not clear whether convection is a necessary ingredient of the explosion itself. Roughly one per cent of the collapse energy released as neutrinos is recaptured to drive the explosion. Why one per cent and not, say, ten per cent? We don't know.

Evolution of the emission line spectrum

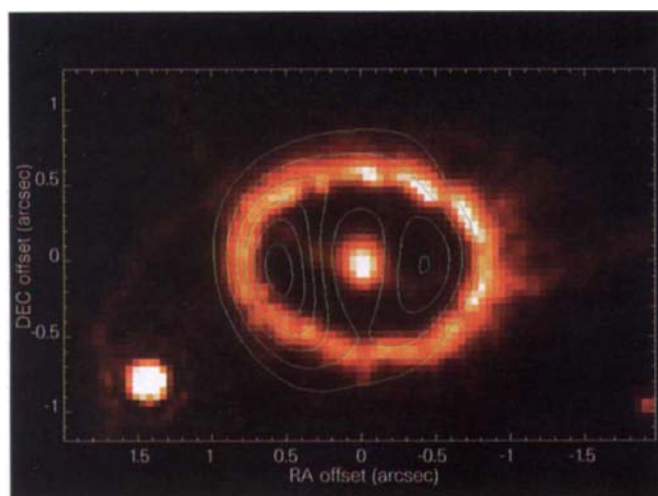


Figure 1 Radio and optical images of SN1987A and its inner ring⁷. The bright radio lobes seen at 9 GHz (contour lines) are synchrotron radiation from relativistic electrons accelerated by the supernova blast wave, which is now nearing the inner circumstellar ring.

has given us a good picture of the structure and thermodynamics of the explosion debris (C. Fransson, Stockholm Obs.). After the explosion, the debris was mixed on large scales by instabilities, but it has remained chemically inhomogeneous. The observed degree of mixing requires that the supernova progenitor was asymmetric, possibly owing to strong convection driven by thermonuclear burning (D. Arnett, Univ. Arizona).

The debris is now very cold throughout. Less than two years after the explosion, the iron-rich and oxygen-rich components cooled by radiation to a few hundred K (the thermal catastrophe), and now even the helium- and hydrogen-rich components have cooled, by adiabatic expansion¹. Today, the optical emission from the supernova debris is dominated by recombination of residual hydrogen ions and atomic excitation resulting from the radioactive decay of about 10^{-4} solar masses of ^{44}Ti .

Why was the progenitor of SN1987A a blue supergiant star rather than a red supergiant, as is thought to be the case for most type-II supernovae? Perhaps, as the binding energy of a massive star is insensitive to the radius of its outer envelope, subtle changes in the opacity or mean molecular weight of the envelope allowed it to shrink, heating it and changing its colour from red to blue (S. Woosley, Univ. Calif. Santa Cruz).

Another model that leads to a blue supergiant is one in which the progenitor formed from the merger of two stars in a binary system². Evidence in favour of the merged binary hypothesis comes from Hubble Space Telescope images of SN1987A. They show a bright ring around the star^{3,4} (Fig. 1), which can be explained as a real physical ring, ejected as an equatorial outflow of several solar masses of gas during a merger of the two stars some 20,000 years before the explosion. Such a merger would probably yield a progenitor that is highly flattened by rotation. If so, the explosion would naturally blow out preferentially along the polar axis, as indicated by recent images showing that the glowing debris of the supernova itself is elongated

perpendicular to the ring plane (J. Pun, NASA/Goddard). But although such a model may be plausible, it is not well developed, and certainly not universally accepted. Moreover, it does not provide a simple explanation for the origin of the outer circumstellar loops. The inner ring and the outer loops are more probably thin, flash-ionized layers at the inner surfaces of a much greater mass of circumstellar matter, as yet unseen.

When SN1987A aficionados converge again to celebrate its 20th anniversary, we should have a good chance to unravel some of its mysteries. The blast wave from the supernova will strike the inner ring some six to ten years from now^{5,6}. When it does, the resulting shock wave will cause the ring to brighten by a factor $\sim 10^3$ in all bands of the electromagnetic spectrum (except γ -rays).

The ionizing radiation from the interaction will illuminate hitherto unseen circumstellar matter. Moreover, the impact will give us an unprecedented experiment in the physics of interstellar shocks, including a chance to observe the acceleration of relativistic electrons in real time.

In fact, we are already seeing precursors of this impact, in non-thermal radio emission seen by the Australia Telescope Compact Array⁷ and X-ray emission seen by ROSAT⁸. Evidently, the supernova blast has already encountered a region of relatively dense ionized hydrogen inside the inner ring⁵. Although we don't yet have a clear picture of this shock, a model for the X-ray emission⁶ suggests that the shocked gas should be bright enough in ultraviolet emission lines to see with the Space Telescope Imaging Spectrometer that was installed during the last Hubble Space Telescope servicing mission. If so, we should soon have images and spectra of sufficient quality to locate the blast wave and tell us when it will strike the ring. □

Richard McCray is in the Department of Astrophysical and Planetary Sciences, University of Colorado, Boulder, Colorado 80309-0440, USA.

1. Wang, L. *et al. Astrophys. J.* **466**, 998–1010 (1996).
2. Podsiadlowski, P. *Publ. Astron. Soc. Pacif.* **104**, 717–729 (1992).
3. Blondin, J. M. & Lundqvist, P. *Astrophys. J.* **405**, 337–352 (1993).
4. Martin, C. L. & Arnett, D. *Astrophys. J.* **447**, 378–390 (1995).
5. Chevalier, R. A. & Dwarakadas, V. V. *Astrophys. J.* **452**, L45–L48 (1995).
6. Borkowski, K., Blondin, J. & McCray, R. *Astrophys. J.* **476**, L31–L34 (1997).
7. Gaensler, B. M. *et al. Astrophys. J.* (in the press).
8. Hasinger, G., Aschenbach, B. & Trümper, J. *Astron. Astrophys.* **312**, L9–L12 (1996).

Mechanics

Patterns of stress in crumpled sheets

Gerhard Gompper

The forced crumpling of thin elastic sheets is experienced frequently in everyday life. Typical examples are the crumpling of paper — which is particularly familiar to scientists — the crumpling of soda cans after use, and the crumpling of metal sheets in car accidents. Despite the enormous importance of this phenomenon, our understanding of the behaviour of strongly deformed elastic sheets is limited. In a paper published on 17 February in *Physical Review Letters*, Kramer and Witten¹ demonstrate for the first time that a large part of the total elastic energy of a crumpled sheet is contained in a network of narrow, stretching ridges. This result could open the door, for example, to the rational design of energy-absorbing materials.

It all began with the attempt by Witten and Li² to understand the shape of very large fullerene balls. In the limit of large size,

deformation can be described by continuum elastic theory. There are two contributions to the elastic energy: the stretching energy, which is the cost of changing the bond length of neighbouring carbon atoms from minimum-energy values; and the bending energy, which describes the energy cost of tilting adjacent carbon rings with respect to one another. If there were no energy cost to bend the surface, the stable shape of the molecule would be a regular, flat-sided icosahedron (with a carbon pentagon at each of the twelve corners) — a shape that requires no bond stretching. When the bending rigidity is finite, however, the sharp edges of the icosahedron are expected to soften so that the minimum-energy surface is smooth except near the corners.

One of the appealing features of this kind of research is that many effects can be seen qualitatively in simple 'paper experi-

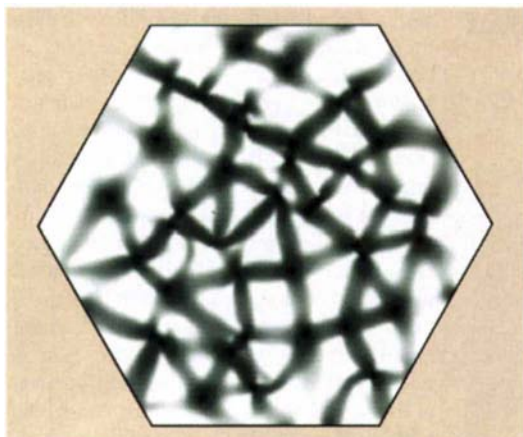


Figure 1 Curvature energy distribution in a hexagonal sheet of diameter L , which has been crushed into a sphere of radius $R_0 \approx L/6$. Darker regions have higher energy density. (From ref. 1.)

ments²⁻⁴. I would therefore like to encourage readers to study the geometry shown in the panel (below) with a sheet of paper. It can be easily seen in such an experiment that the curvature is localized near the ridge that connects two opposing corners.

Witten and Li² give a very simple argument for the shape of the edge connecting two corners, and the dependence of edge energy on scale (see panel). The simplicity of their argument means that it has had to be checked carefully to test its validity and its range of applicability. This has been done by computer simulations of simple models of triangulated surfaces, which were introduced some time ago by Kantor and Nelson³. The numerical studies, carried out by Witten and co-workers⁴ and by Kroll and co-workers⁶, show that the scaling laws do indeed describe the asymptotic behaviour of large fullerene balls. However, quite large system sizes are necessary in order to observe this

behaviour, with edge lengths 1,000 times the effective thickness of the sheet or more. For smaller systems, such as C_{60} , the bending energy of the cone-shaped regions near the corners dominates^{4,6}.

Kramer and Witten¹ have now gone beyond the study of isolated stretching ridges. In their computer simulation, a nearly flat sheet is approximated by a triangular network of springs with some bending elasticity. They put the sheet into a spherical shell, and slowly decrease the shell radius R_0 until it is much smaller than the diameter L of the initial sheet. The resulting distribution of the curvature energy is shown in Fig. 1. It clearly demonstrates the formation of stretching ridges. About 40% of the energy is localized in very small areas (vertices), which correspond to the corners of the fullerene balls; the next 40% is contained in the narrow ridges that connect the vertices. Since the length of each ridge is found to be

similar to the radius R_0 of the confining sphere, their number must go as $(L/R_0)^2$, so the total elastic energy should scale as $R_0^{-5/3}$ (see panel). The simulation results are consistent with this prediction.

The theory of crumpled sheets applies to macroscopic as well as to microscopic elastic sheets. The work of Witten and co-workers^{1,2,4}, which led to the present understanding of the crumpled state, originated in studies of the properties of fluid and polymer membranes^{7,8}. For these microscopic surfaces, thermal fluctuations can be important. It has been shown, for example, that thermal fluctuations crumple non-self-avoiding polymerized membranes with small bending rigidities⁵ without any external compression. And fluid membranes, which have no stretching energy, collapse into crumpled, branched-polymer-like shapes for sufficiently low bending rigidity due to thermal fluctuations⁹, even with self-avoidance.

But self-avoidance stabilizes the flat phase¹⁰ in all real polymerized membranes, such as graphite-oxide sheets¹¹ or the spectrin network of red blood cells¹², giving them the elastic stiffness relevant to the new study. Further, thermal fluctuations should not modify the scaling behaviour of stretching ridges¹³. So knowledge about ridges may help us understand the passage of red blood cells through narrow capillaries¹⁴, for example.

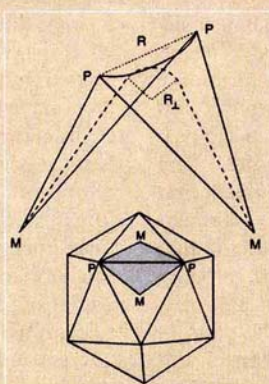
It is the subtle interplay between stretching, bending and thermal fluctuations which makes this field so exciting, and which gives microscopic membranes their unique properties. The work of Kramer and Witten is a beautiful example of the contribution that physics can make to materials research and mechanical engineering, and it should have important implications for the understanding of biological systems. □

Gerhard Gompper is in the Max-Planck-Institut für Kolloid- und Grenzflächenforschung, Kantstrasse 55, 14513 Teltow, Germany.

1. Kramer, E. M. & Witten, T. A. *Phys. Rev. Lett.* **78**, 1303–1306 (1997).
2. Witten, T. A. & Li, H. *Europhys. Lett.* **23**, 51–55 (1993).
3. Kantor, Y., Kardar, M. & Nelson, D. R. *Phys. Rev. A* **35**, 3056–3071 (1987).
4. Lobkovsky, A., Gentges, S., Li, H., Morse, D. & Witten, T. A. *Science* **270**, 1482–1485 (1995).
5. Kantor, Y. & Nelson, D. R. *Phys. Rev. Lett.* **58**, 2774–2777 (1987); *Phys. Rev. A* **36**, 4020–4032 (1987).
6. Zhang, Z., Davis, H. T., Maier, R. S. & Kroll, D. M. *Phys. Rev. B* **52**, 5404–5413 (1995).
7. Nelson, D., Piran, T. & Weinberg, S. (eds) *Statistical Mechanics of Membranes and Surfaces* (World Scientific, Singapore, 1989).
8. Lipowsky, R. *Nature* **349**, 475–481 (1991).
9. Gompper, G. & Kroll, D. M. *Phys. Rev. Lett.* **73**, 2139–2142 (1994); *Phys. Rev. E* **51**, 514–525 (1995).
10. Abraham, F. F., Rudge, W. E. & Plischke, M. *Phys. Rev. Lett.* **62**, 1757–1759 (1989).
11. Spector, M. S., Naranjo, E., Chiruvolu, S. & Zasadzinski, J. A. *Phys. Rev. Lett.* **73**, 2867–2870 (1994).
12. Schmidt, C. F. et al. *Science* **259**, 952–955 (1993).
13. Zhang, Z., Davis, H. T. & Kroll, D. M. *Z. Phys. B* **97**, 337–340 (1995).
14. Reinhart, W. H. et al. *Biorheology* **28**, 537–549 (1991).

Scaling ridges

Witten and Li² give a very simple scaling argument for the shape of the edge connecting two corners of an elastic icosahedron, depicted in the figure here. Any deformation that involves most of the bonds on the surface costs an energy proportional to the area; therefore, stretching has to be confined to narrow ridges along the edges. The surface near the edge is now assumed to have a roughly cylindrical shape, with a radius of curvature R_1 at its midpoint. For a distance R between the corners, the total curvature energy is approximately $\kappa A R_1^{-2}$,



where κ is the bending rigidity and $A \approx \pi R R_1$ is the area of the curved region. The curvature of the edge implies that the mid-line of the bend retracts inward. Because the distance R cannot change (this would require the faces of the icosahedron to be

compressed, at a prohibitively large cost in stretching energy), the length of the mid-line must therefore increase by a fraction of order $\gamma = (R_1/R)^2$. The fraction γ is proportional to the change of the carbon-carbon bond length along the ridge, so the bond stretching energy is about $K_0 A \gamma^2$, where K_0 is the bond's elastic modulus. Minimization of the total energy then gives the result that the radius of the edge curvature scales as $R_1 \approx R^{2/3}$, and the total edge energy as $\sim R^{1/3}$. So for large R , the curvature is indeed concentrated into narrow (small R_1) ridges. GG

Another opiate for the masses?

David Julius

The active ingredients of natural toxins and herbal remedies have helped to unlock the secrets of biological processes ranging from cell division to synaptic transmission. A notable example is opium — an extract of the poppy plant that has been used for centuries to induce euphoria and relieve pain. The active component of opium is morphine, and its synthetic congener is heroin^{1,2}. The analgesia, reward-seeking behaviour and physical dependence elicited by morphine^{2,3} are mainly mediated through the μ opioid receptor and, on page 499 of this issue, Zadina and colleagues⁴ report the isolation of two new μ -selective peptide ligands from mammalian brain. These peptides (called endomorphin-1 and -2) rival morphine in both their potency as agonists and their analgesic activity. If they are, indeed, natural ligands for the μ receptor, then they should provide new molecular tools for dissecting the endogenous opiate pathways in the brain and spinal cord.

The quest to determine how the opiate alkaloids mediate euphoric, analgesic and addictive states has helped to shape a central tenet of modern neuropharmacology: namely, that toxins and drugs usurp sites at which endogenous factors normally act to transduce biological signals. Two discoveries affirmed this hypothesis. First, radiolabelled opiate alkaloids were shown to bind to specific, high-affinity sites in the brain, corresponding to three distinct guanine-nucleotide-binding (G) protein-coupled opioid-receptor subtypes, termed κ , μ and δ ⁵ (Fig. 1). Second, endogenous morphine-like substances were identified in the brain. These so-called endorphins are peptides that activate opioid-receptor subtypes, with characteristic rank-order potencies. But, until now, no single endorphin had been shown to bind with both high selectivity and high affinity to the μ receptor, begging the question of whether the true 'endogenous morphine' had, in fact, been discovered.

The route that was taken by Zadina *et al.*⁴ to uncover the new endomorphins is interesting in itself, because it involved an unusual interplay between pharmacology and combinatorial chemistry. The story begins with the isolation of a tetrapeptide called Tyr-MIF-1 (Tyr-Pro-Leu-Gly-NH₂) from bovine and human brain⁶. This peptide acts as both an agonist and an antagonist in physiological assays for binding to the μ receptor, which measure the opiate-dependent inhibition of electrically induced contractions in the guinea-pig ileum. An antibody was generated to Tyr-MIF-1, then a radio-immunoassay was used to isolate a related

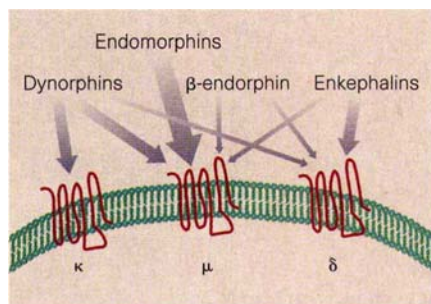


Figure 1 Known agonists for the κ , μ and δ opioid receptors. Zadina *et al.*⁴ have identified two naturally occurring endorphins (termed endomorphin-1 and -2) which bind with high specificity and high affinity to the μ receptor. Relative affinities of various agonists for different opioid-receptor subtypes are depicted by arrows.

peptide from mammalian brain extracts⁷, Tyr-W-MIF-1 (Tyr-Pro-Trp-Gly-NH₂). Tyr-W-MIF-1 is a more potent opiate agonist than Tyr-MIF-1, and it has a respectable selectivity for μ receptors over δ or κ receptors (200–300-fold). But its affinity for μ receptors (based on competitive binding assays with the μ -selective endorphin analogue ³H-DAMGO) is modest (K_i 70 nM).

To identify a peptide with greater affinity for the μ receptor, Zadina *et al.* used a solid-phase synthesis technique known as polyethylene pin technology to generate 20 Tyr-W-MIF-1 derivatives, representing all of the possible amino-acid substitutions at the fourth position of the peptide. Phe4 — the peptide with phenylalanine at the fourth site (Tyr-Pro-Trp-Phe-NH₂) — turned out to have the desired characteristics, including sub-nanomolar affinity for μ receptors, and very high selectivity of binding to μ versus δ (>4,000-fold) or κ (>15,000-fold) sites. When it was injected into the brain or spinal column of mice, synthetic Phe4 proved to be as potent as morphine in producing a long-lasting analgesic response. Moreover, the response was reversed by administration of broad-spectrum or μ -selective opiate antagonists.

Next, the authors took a daring leap of faith — they predicted that the nervous system might express its own version of the Phe4 peptide. Using an antibody that was specific for the Phe4 peptide, they isolated immunologically crossreactive material from bovine brain extracts. These extracts indeed contained endogenous Phe4 peptide (endomorphin-1), along with the closely related species Tyr-Pro-Phe-Phe-NH₂ (endomorphin-2). Because other endorphins contain the amino-terminal signature sequence Tyr-Gly-Gly-Phe, it may be that a stable structure imposed by the proline residue at the second position helps to

increase selectivity for the μ receptor. Furthermore, the long-lasting analgesic activity of endomorphin-1 and -2 could reflect protection from exoproteolytic degradation conferred by the carboxy-terminal amidation.

By analogy with other peptide hormones and neurotransmitters⁸, including the endorphins, one would expect the endomorphins to be excised from a larger polypeptide precursor. Cloning the gene for this hypothetical pro-endomorphin would serve at least three purposes. First, it would confirm that the endomorphins are indeed endogenous neuropeptides, derived by specific proteolytic cleavage events. Second, sequence analysis of the gene would clarify any genetic relationship between endomorphins and the peptides that led to their isolation, including Tyr-MIF-1 and Tyr-W-MIF-1, and other biologically active peptides might be discovered in the process. Third, the endomorphin gene would provide sensitive and specific molecular probes for determining where these peptides are expressed in the nervous system. If the endomorphins are the endogenous ligands for μ receptors, they might be found in pain-modulating, reward-seeking or emetic centres of the nervous system (for example, the spinal-cord dorsal horn, periaqueductal grey, area postrema, striatum, amygdala and nucleus accumbens)². Nonetheless, questions about the peptide distribution can perhaps be addressed with the specific antibody that was used by Zadina and colleagues⁴ to purify the endomorphins.

Could the endomorphins provide relief from pain without eliciting the negative symptoms that are associated with morphine, such as respiratory depression, nausea, tolerance and addiction? Although pharmacological and genetic studies³ indicate that this is unlikely, the answer may be 'yes' if the endomorphins are more μ -selective than morphine, or if the activation of μ receptors by structurally distinct agonists has differential effects on signalling. We next need to examine interactions between the endomorphins and cell lines that express cloned opioid receptors. Finally, does this work signal the beginning or the end of the search for endogenous opiate peptides? The modest binding selectivity of other known endorphins to the three opioid-receptor subtypes indicates that peptides with comparable selectivity for the δ and κ receptors are out there waiting to be discovered. □

David Julius is in the Department of Cellular and Molecular Pharmacology, University of California at San Francisco, San Francisco, California 94143-0450, USA.

1. Snyder, S. *Drugs and the Brain* (Freeman, New York, 1986).
2. Fields, H. *Pain* (McGraw-Hill, New York, 1987).
3. Matthes, H. *et al.* *Nature* **383**, 819–823 (1996).
4. Zadina, J. E., Hackler, L., Ge, L.-J. & Kastin, A. J. *Nature* **386**, 499–502 (1997).
5. Kieffer, B. *Cell. Mol. Neurobiol.* **15**, 615–635 (1995).
6. Horvath, A. & Kastin, A. J. *Biol. Chem.* **264**, 2175–2179 (1989).
7. Erchegyi, J., Kastin, A. & Zadina, J. *Peptides* **13**, 623–631 (1992).
8. Herbert, E. & Uhler, M. *Cell* **30**, 1–2 (1982).

The origin of maize branches out

Rob Martienssen

The origins of cultivated maize have puzzled evolutionary biologists for many years. The upright, unbranched maize plant with its large, closely packed ear, can survive only by careful cultivation and propagation, and it seems implausible that maize could be derived from its wild grass-like relatives: how could such rapid evolution have occurred during the relatively recent course of agricultural history? On page 485 of this issue, Doebley *et al.*¹ go some way to resolving this question — they describe a molecular component of one of the traits that distinguishes maize from its closest wild relative, teosinte.

Almost 60 years ago, Mangelsdorf and Reeves² found that, when maize was crossed with teosinte, segregation of just four or five genetic factors could explain the frequency of offspring that resembled the parents in later generations. They suggested that these factors arose through the exchange of large chromosomal regions between hypothetical wild relatives of maize and the related genus *Tripsacum*^{2,3}. Mangelsdorf argued against the alternative view — that maize was derived from teosinte by descent — on the grounds that the rate of spontaneous evolution was far too slow to explain the large morphological changes that were involved. Then, a year later, George Beadle⁴ proposed that each of the four or five chromosomal regions represented an individual gene (one gene, one trait⁵). He argued that maize could have been selected as a new variant of teosinte, separated by a handful of major mutations, and that early farmers might have propagated teosinte for its value as a somewhat explosive popcorn⁴.

The origin of maize has been hotly debated ever since. Ilitis⁶, for example, agreed with Beadle that maize was derived directly, and rather suddenly, from teosinte. Moreover, Ilitis realized that alteration of plant architecture (specifically, the pattern of branching) was a crucial parameter in the transition from teosinte to maize. But he found genetic explanations wanting, and instead suggested that there had been an epigenetic sexual catastrophe. He proposed that the male inflorescence, which tips the axillary branches of teosinte, underwent a sex-change that led, through changes in nutrient allocation and other environmental influences, to the maize ear as we know it.

More recently, the genetic differences between maize and teosinte have been revisited using analysis of quantitative trait loci (QTL). QTL are genetic loci that contribute to multifactorial traits and can be mapped simultaneously using molecular markers. By choosing appropriate traits, Doebley and co-workers⁷ mapped five major loci that were

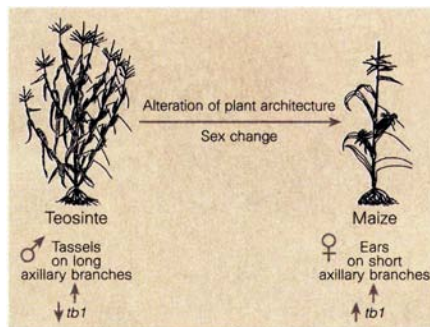


Figure 1 The argument that cultivated maize is descended from its closest wild relative, teosinte, is strengthened by new findings from Doebley *et al.*¹. They have cloned the teosinte branched1 (*tb1*) gene, which is responsible for one of the main differences between maize and teosinte — apical dominance. Whereas teosinte plants bear many long axillary branches that are tipped by male inflorescences (tassels), maize plants have fewer, short axillary branches tipped by female ears. The *tb1* gene controls this transition and the growth of reproductive organs, which may be further modified by the terminal ear1 gene. Moreover, maize *tb1* is expressed at much higher levels than teosinte *tb1*, indicating that maize may have diverged from teosinte due to changes in genetic regulation.

responsible for the difference between teosinte and maize, along with a number of minor modifiers. Following Robertson's suggestion that QTL can correspond to individual genes, Doebley *et al.*⁷ looked for candidate genes that mapped to the five QTL in maize. One of these genes was teosinte branched1 (*tb1*), mutations in which lead to the outgrowth of basal lateral branches as tillers and upper lateral branches as shoots, replacing the familiar maize ear.

Allelism tests showed that, in maize, *tb1* corresponds to the QTL for the growth pattern of teosinte, although interactions with

other loci also contribute to the overall architecture⁷. For example, when the maize *tb1* locus was crossed into teosinte, short branches tipped with female teosinte ears replaced the long branches tipped with male tassels. This sexual transformation was greatly enhanced by a major QTL on chromosome 3, but the effects of this QTL alone were almost undetectable in teosinte. A candidate for this QTL is the maize terminal ear1 (*te1*) gene, which maps to this region⁷.

The latest chapter in this story is the molecular characterization of the *tb1* locus, reported by Doebley *et al.*¹. Transposon tagging was used to identify several alleles of *tb1*, all of which mapped to a single transcriptional unit. They found that *tb1* encodes a protein with a novel domain that is also found in the snapdragon *cycloidea* gene⁸, mutations in which result in radially, rather than bilaterally, symmetrical flowers.

Although the significance of the similarity between *cycloidea* and *tb1* is not immediately apparent, there are some intriguing parallels. Both genes suppress the growth of certain floral organs (although more detailed investigations of the *tb1* expression patterns are needed). Perhaps more importantly, *tb1* and *cycloidea* also organize axillary structures — branches and flowers, respectively — relative to the main axis^{1,8}. Expression studies indicate that *cycloidea* acts at a distance⁸, almost like an organizing centre. So perhaps the epistatic interactions⁷ between *tb1*, and the QTL that affects the architecture of the main inflorescence (*te1*?), have a similar basis in maize.

Many of the mutations that distinguish maize and teosinte show some degree of dominance. This may reflect the irreversible nature of loss-of-function mutants during evolution⁷ or, alternatively, it may reflect redundancy of the maize and teosinte genomes, which have undergone extensive chromosomal duplications⁹. Genetic evidence indicates that the *tb1* gene of cultivated maize behaves like a hypermorph (a dominant mutation) when it is backcrossed into teosinte⁷. Molecular analysis bears this out, as Doebley *et al.*¹ find that the teosinte *tb1* gene has only minor amino-acid changes compared with the *tb1* gene in maize,



Figure 2 Evolving side by side — teosinte growing with maize on a hillside in Guerrero, Mexico. (Figure courtesy of J. Doebley.)

but it is expressed at considerably lower levels. The altered expression of *tb1* in maize seems to suppress second-order branching and promote condensed ear-shoot morphology. Such expression patterns may sometimes occur in teosinte as a result of environmental stress, so conferring a less extravagant, unbranched stature^{1,6,7}.

The conversion of teosinte to maize represents a change in plant architecture of historic proportions, and *tb1* seems to have been crucial to this transition. Isolation of the other key genes is under way, and we can hope to see a molecular explanation. As comparative analysis of plant genomes proceeds, genes that have major effects on plant architecture, such as the *cauliflower* gene of *Arabidopsis*¹⁰ or the *tb1* gene of maize, are being recognized for their role in the origin of crop plants. With the genes responsible to hand, it could be possible to manipulate crop plants to have radically altered properties.

Most commentaries on plant biotechnology typically focus on plant disease resistance, food quality or agrochemicals, as the main benefits of plant genetic engineering. But plant biotechnologists would do well to take a lesson from Aztec farmers, and engineer plants that have completely new forms for a new generation of agriculture. □

Rob Martienssen is at Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724, USA.

1. Doebley, J., Stec, A. & Hubbard, L. *Nature* **386**, 485–488 (1997).
2. Mangelsdorf, P. C. & Reeves, R. G. *Proc. Natl Acad. Sci. USA* **24**, 303–312 (1938).
3. Mangelsdorf, P. C. & Reeves, R. G. *J. Hered.* **22**, 329–343 (1931).
4. Beadle, G. W. *J. Hered.* **30**, 245–247 (1939).
5. Dorweiler, J., Stec, A., Kermicle, J. & Doebley, J. *Science* **262**, 233–235 (1993).
6. Iltis, H. *Science* **222**, 886–894 (1983).
7. Doebley, J., Stec, A. & Gustus, C. *Genetics* **141**, 333–346 (1995).
8. Luo, D., Carpenter, R., Vincent, C., Copsey, L. & Coen, E. *Nature* **383**, 794–800 (1996).
9. Helentjaris, T., Weber, D. & Wright, S. *Genetics* **118**, 353 (1988).
10. Kempin, S. A., Savidge, B. & Yanofsky, M. F. *Science* **267**, 522–525 (1995).

Luminous infrared galaxies

Diagnostics of the power supply

Willem A. Baan

Astronomical masers are observed in star-forming regions or stellar objects in our Galaxy, but also in the nuclear regions of other galaxies. Certain molecules can be 'pumped' into an excited state by their environment, such that they amplify their own line emissions or some background continuum radiation. This 'avalanche of photons' results in luminous maser emission lines that serve as valuable diagnostics of the physical environment.

Strong interstellar OH (hydroxyl) masers in our Galaxy are mostly pumped by collisional processes and shocks, but OH 'megamasers' in the nuclei of other galaxies are typically a million times more luminous. The report announcing the discovery of the OH megamaser in the interacting galaxy system Arp220 (or IC4553 as it was called before becoming famous) attributed the pumping of the OH masers to the far-infrared (FIR) radiation field¹. A paper by Skinner *et al.* (page 472 of this issue²) provides the long-awaited confirmation that the driving mechanism for OH-megamaser activity is indeed the strong FIR radiation field within these galaxies. Moreover, these new data, obtained with the Infrared Space Observatory (ISO), bear on the much debated nature of the underlying energy source for these emissions.

The discovery of OH megamasers came simultaneously with the recognition of a new class of (super)luminous infrared galaxies within the IRAS (Infrared Astronomy Satellite) database. The FIR emission of these galaxies is more or less isotropic, while the OH-megamaser emission is likely to be

beamed towards us. The OH-megamaser sources are thus a representative subsample of all luminous FIR galaxies. However, the energy source of these FIR galaxies and their proper place in schemes of galaxy evolution and formation have not been resolved. There are two views on the issue. It could be that the energy source at the Galactic Centre is an active galactic nucleus (AGN), with a massive black hole; alternatively, the source could be a powerful nuclear starburst. Both would be hidden deep within a thick cocoon of dust that re-radiates much of the original energy at infrared wavelengths.

Optical surveys of these luminous FIR sources suggest that the fraction of AGNs among them increases with the FIR luminosity, as does the fraction of interacting, colliding or merging systems. During a merger, large amounts of gas are liberated and deposited in the central regions of the galaxy. Through accretion flows, gas deposited in the nuclear region could activate or reactivate a dormant nucleus, while at the same time the strong interaction could also trigger bursts of star formation in the region. On the other hand, the radio characteristics of these FIR galaxies suggest that most of the radio emission comes from starbursts^{3,4}.

There is no strong evidence for dominant radio AGNs. This does not mean that starbursts and weak AGNs cannot co-exist. As many as 50 per cent of the luminous FIR galaxies may have weak (a few per cent of total flux), compact AGN-like radio sources embedded in the starburst emissions^{5,6}. Similarly, recent studies of sources in the optical transition region between starburst and



100 YEARS AGO

Prof. Schiaparelli has just published an account of his observations of Mars made during the years 1883–84. These observations are contained in the publication of the *Reale Accademia dei Lincei*, 1895–96, and are accompanied by a new chart of the surface-markings and eight diagrams of the planet's disc. Of the thirty-one canals seen to develop in 1881 and 1882, only eighteen have been observed again at this opposition, while seven new ones have been noted. In the polar gap a large dark rift was seen, apparently separating the snow into two unequal portions. An examination of the chart with that published by Lowell will be found to give many interesting points of difference. In the March number of the *Bulletin de la Société Astronomique de France* there will be found a brief account of these observations, with reproductions of the plates in Schiaparelli's original memoir. There are added a series of drawings of the planet observations, made from the Observatory of Juvisy, at the opposition in December last. — "The Planet Mars." From *Nature* 1 April 1897.

50 YEARS AGO

The Section of Psychology, lately formed inside the Association of Scientific Workers, met for the first time at Bedford College, London, on January 18. ... The primary responsibility of psychologists as such is to work whole-heartedly for the advancement of their science. To-day in most sciences comparatively little can be done by individual workers, and the importance of co-ordinated research here as elsewhere was stressed. The time may even have come for university departments of psychology to establish a co-ordinating body so that (1) research workers can know what others are doing at present in their field and thus avoid duplication, and (2) plans for research into fundamental problems can be developed. An urgent need is to let the public — at present badly instructed — know what psychology can and cannot do. More use could be made of the ordinary media of communication by psychologists who can express complicated ideas simply and clearly; but this might require a changed attitude upon the part of some editors.

From *Nature* 5 April 1947.

AGN show that both are present in some sources^{7,8}. Optically, the two merging nuclei of Arp 220 look like Seyfert 2 (AGN) sources, while at radio wavelengths the nuclei look very much like starbursts. This is also true for several of the other OH megamasers.

Other ISO results shed light on this controversy. Arp220 and other sources do not exhibit high-ionization spectral lines in the infrared spectrum, as would have been the trademark of an underlying AGN (refs 9,10). Furthermore, the large amounts of dust almost completely absorb any optical radiation coming from the nuclear region itself, while the strong FIR lines partially pass through. Modelling of the bursts of star formation suggests that these bursts are relatively recent and that they completely dominate the production of nuclear energy.

Skinner *et al.*² add independent arguments to this discussion. Their data show a strong absorption line in the FIR spectrum at 35 μm , which corresponds to one of the crucial pumping transitions of the OH molecule. On the basis of this line, the authors conclude that the infrared emission region in Arp220 is unlikely to be a point source, but is probably a more extended cloud, possibly along the lines of a dusty torus. The optical characteristics of obscured FIR nuclei may thus be rather misleading, and it now seems increasingly reasonable that what appears to be an AGN may in fact be a highly obscured starburst.

Megamasers have been found in four molecular flavours: hydroxyl (OH), methylidyne (CH), formaldehyde (H_2CO) and water vapour (H_2O). All of these emissions are likely to be beamed at us and provide a diagnostic picture of the molecular material along the line of sight towards the nucleus of the FIR galaxy. The H_2O megamasers tell us about molecular material in a very compact disk buried deep inside the nuclear region of galaxies with less dust and less FIR emission, but with an AGN (black hole) at the centre. On the other hand, the OH megamasers tell us about the dusty molecular torus in a whole class of luminous FIR galaxies harbouring starbursts and maybe weak AGNs.

Willem A. Baan is at the Arecibo Observatory of the National Astronomy and Ionosphere Center, Barrio Esperanza, Arecibo, Puerto Rico 00613.

1. Baan, W. A., Wood, P. A. D. & Haschick, A. D. *Astrophys. J.* **260**, L49–L52 (1982).
2. Skinner, C. J. *et al. Nature* **386**, 472–474 (1997).
3. Condon, J. J. *et al. Astrophys. J. Suppl.* **73**, 359–400 (1990); *Astrophys. J.* **378**, 65–76 (1991).
4. Lehnert, M. D. & Heckman, T. M. *Astrophys. J.* **472**, 546 (1996).
5. Norris, R. P., Allen, D. A., Sramek, R. A., Kesteven, M. J. & Troup, E. R. *Astrophys. J.* **359**, 291–295 (1990).
6. Lonsdale, C. J., Smith, H. E. & Lonsdale, C. J. *Astrophys. J.* **405**, L9–L12 (1993).
7. Ashby, M. L. N., Houck, J. R. & Matthews, K. *Astrophys. J.* **447**, 545 (1995).
8. Veron, P., Gonçalves, A. C. & Veron-Cetty, M.-P. *Astron. Astrophys.* (in the press).
9. Sturm, E. *et al. Astron. Astrophys.* **315**, L133–L136 (1996).
10. Lutz, D. *et al. Astron. Astrophys.* **315**, L137–L140 (1996).

Immunology

Viral decoy vetoes killer cell

Klas Kärre and Raymond M. Welsh

In his book *River out of Eden*¹, Richard Dawkins discusses how birds protect their nest and offspring from predators. Citing the Austrian zoologist Wolfgang Schliedt, he points out that the “rule of thumb a turkey mother uses to recognize nest robbers is a disarmingly brusque one: in the vicinity of the nest, attack anything that moves, *unless* it makes a noise like a baby turkey”. Clearly, a shrewd intruder should avoid contact with the aggressive mother turkey, either by making himself invisible (‘stealth’ strategy) or by imitating the cry of baby turkeys (‘decoy’ strategy). Perhaps the wisest would be to adopt both types of strategy, and this is just what cytomegalovirus (CMV) seems to have done to avoid aggressive contacts with killer cells of the immune system.

Examples of the stealth approach — where CMV and other viruses interfere with expression of the major histocompatibility complex (MHC), which is normally used for cellular antigen presentation and immune surveillance — have already been described^{2,3}. Now, on pages 510 and 514 of this issue, respectively, Farrell *et al.*⁴ and Reyburn *et al.*⁵ provide evidence for the existence of a decoy strategy. They show that the products of homologous genes in murine⁴ and human⁵ CMV protect virus-infected cells from natural killer (NK) cells, presumably by imitating normal MHC molecules. This decoy misleads the NK cells, which normally attack anything that looks like a cell, *unless* it expresses self-MHC molecules as a veto signal^{6,7}.

There are two classes of MHC molecule, allowing for continuous quality control of a host cell’s external and internal protein environments. Class I molecules bind small peptide fragments that are produced within the cell, and transport them to the surface for scrutiny by T lymphocytes (Fig. 1). These T lymphocytes dock onto peptide-charged MHC molecules through specific T-cell receptors, but they only react when the MHC molecules carry non-self peptides. This results in the development of aggressive cytotoxic T lymphocytes, which limit viral replication and spread by secreting cytokines or by directly killing infected cells that carry the non-self peptides that initiated the reaction.

Although sophisticated, this system is vulnerable. Imagine a virus (or any micro-organism) that evolved, or captured, genes that interfere with the class I MHC pathway — the infected cell would then be invisible to T cells. This stealth strategy has indeed been adopted by many viruses^{2,3}; for example, human CMV has at least four gene products that act in concert to reduce the expression of class I molecules. But the mammalian

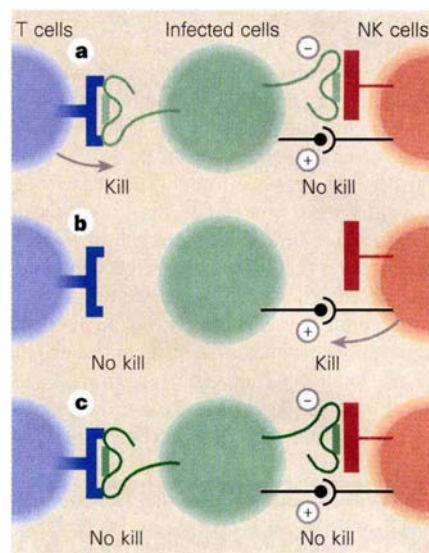


Figure 1 a, Virally infected cells are normally killed by the immune system’s T cells, which recognize viral peptides that are presented by class I molecules of the major histocompatibility complex (MHC; light green). b, Some viruses evade this immune surveillance using a ‘stealth’ strategy — they interfere with expression of the MHC to avoid attack by T cells. However, they are then susceptible to the natural killer (NK) cells, which kill cells that do not express class I MHC molecules. c, Farrell *et al.*⁴ and Reyburn *et al.*⁵ now provide evidence that cytomegalovirus (CMV) also uses a ‘decoy’ strategy. Murine⁴ and human⁵ CMV both express a homologue of class I MHC molecules (dark green), which can act as a decoy to veto the NK cells.

immune system, which has coevolved with these microorganisms, may have at least one countermeasure: the NK cells seem to use a similar strategy to the mother turkey, attacking any cell, unless it expresses class I MHC (ref. 6).

NK cells are equipped with adhesion molecules, triggering receptors and inhibitory receptors⁷. The first two tell the NK cell that it has encountered a potential ‘target’ cell, and they start a series of events that will eventually lead to the destruction of this cell — unless the appropriate MHC class I molecules are detected by inhibitory receptors. The NK cells use different types of inhibitory receptor: some belong to the immunoglobulin-gene superfamily (predominant in humans), whereas others are members of the C-type lectin gene superfamily (predominant in the mouse). Although the receptor families are structurally different, each receptor recognizes a group of MHC class I molecules, and they all mediate inhibitory signals through shared motifs in the cytoplasmic tail.

There are clinical correlations between low NK-cell activity and severe CMV infection⁸, and there are substantial experimental data to indicate that murine CMV is controlled by NK cells⁹. By what mechanism could NK-cell-mediated regulation of CMV infection occur? The answer may lie in the attempts of CMV to escape the cytotoxic T lymphocytes by interfering with the class I pathway for antigen presentation. By removing the 'protective' class I inhibitory structures from the cell surface, the viruses may render the infected cells more sensitive to NK cells. However, both human¹⁰ and murine¹¹ CMV contain a gene that resembles MHC class I genes. The protein products are mainly similar in the extracellular domains, two of which form the groove that accommodates peptide, and the human CMV class I homologue has been shown to engage and present peptides¹². And now, Farrell *et al.*⁴ and Reyburn *et al.*⁵ show that both human and murine viruses use the class I homologues to escape surveillance by NK cells.

Farrell *et al.* constructed a recombinant murine CMV with a functional deletion (designated m144) in its class I homologue. The deletion did not affect replication of murine CMV in fibroblasts *in vitro*, but when mice were infected with the recombinant virus, it was severely attenuated in the spleen and liver. Depletion of NK cells in these mice led to increased synthesis of the recombinant virus, indicating that expression of the class I homologue rendered the virus (at least partially) resistant to NK cells.

The exact mechanism remains to be investigated, but the paper by Reyburn *et al.* sheds some light here. When the human CMV class I homologue (designated UL18) was cloned and expressed in a human cell line that lacked class I MHC, the cells became resistant to NK-cell-mediated lysis. Moreover, UL18 conferred resistance to all NK cells, not just to certain NK clones. This is because the UL18 protein seems to interact with CD94, a component of a heterodimeric inhibitory receptor of the C-type lectin family, which is found on most NK cells and many T cells¹³. So by encoding its own class I homologue, the virus may provide itself with a mechanism to veto NK-cell attack, while using other genes to render the infected cells invisible to T-cell attack.

How does this fit in with what we know about the sensitivity of CMV to NK cells? Fibroblasts that are infected with human CMV *in vitro* are (in contrast to murine CMV-infected cells) quite sensitive to NK-cell-mediated lysis, and various NK-cell clones differ in their ability to lyse infected fibroblasts¹⁴ — seemingly discordant with the results of Reyburn *et al.*⁵. However, expression of UL18 may be tissue specific, and replication in virally infected cells has yet to be detected. So the experiments on the lysis of human CMV-infected cells may have been examining an infection in which the inhibitory UL18 protein was not expressed¹⁴. It would be inter-

esting to see whether cells expressing UL18 might be rendered sensitive to NK cells after infection with human CMV, as they could be displaying both negative and positive signalling structures.

Does murine CMV escape detection by the immune system in a similar manner? NK-cell-mediated control of infection is normal in mice with impaired MHC class I expression due to disruption of the β_2 -microglobulin gene¹⁵, indicating that host-cell class I MHC recognition may not be involved in the NK-cell response to murine CMV. But this point becomes moot if the virus expresses its own class I homologue. Expression of inhibitory receptors on the NK cells, which recognize the class I molecules, is augmented in the β_2 -microglobulin-deficient mice¹⁶. Notably, certain mouse strains can strongly resist infection by murine CMV, with the help of a gene termed *Cmv1*, which maps within (or close to) the gene family that encodes the NK-cell inhibitory receptors for class I molecules¹⁷.

Future studies will probably search for the expression and function of the class I homologue in infected cells *in vivo*. An intriguing question concerns its ability to bind peptides¹². Is there no risk for presentation of viral antigens? Or is this yet another vicious trick to mislead the immune system? Moreover, could host cells evolve countermeasures to lower the expression or sensitivity of the inhibitory receptors so that NK cells cannot 'hear' the veto signal¹⁸? More details are needed before we can conclude whether, and how, CMV behaves like an invisible, but noisy, baby-turkey imitator, and whether the tactic is absolutely safe. Perhaps Dawkins has a clue when he sums up by citing another observation from Schliedt of "a turkey mother that savagely killed all her babies. The reason was woefully simple: she was deaf". □

Klas Kärre is in the Microbiology and Tumor Biology Center, Karolinska Institute, 171 77 Stockholm, Sweden. Raymond M. Welsh is in the Department of Pathology, University of Massachusetts Medical Center, 55 Lake Avenue North, Worcester, Massachusetts 01655-0125, USA.

1. Dawkins, R. *River Out of Eden* (Weidenfeld & Nicolson, London, 1995).
2. Bonifacino, J. S. *Nature* **384**, 405–406 (1996).
3. Spriggs, M. K. *Annu. Rev. Immunol.* **14**, 101–130 (1996).
4. Farrell, H. E. *et al. Nature* **386**, 510–514 (1997).
5. Reyburn, H. T. *et al. Nature* **386**, 514–517 (1997).
6. Kärre, K. *et al. Nature* **319**, 675–678 (1986).
7. Parham, P. (ed.) *Immunol. Rev.* **155**, 5–221 (1997).
8. Biron, C. A., Byron, K. S. & Sullivan, J. S. *New Engl. J. Med.* **320**, 1731–1735 (1989).
9. Tay, C. H. & Welsh, R. M. *J. Virol.* **71**, 267–275 (1997).
10. Beck, S. & Barell, B. B. *Nature* **331**, 269–272 (1988).
11. Rawlinson, W. D., Farrell, H. E. & Barrell, B. G. *J. Virol.* **70**, 8833–8849 (1996).
12. Fahnestock, M. *et al. Immunity* **3**, 583–590 (1995).
13. Aramburu, J. *et al. J. Immunol.* **144**, 3238–3247 (1990).
14. Mason, P. D., Sissons, J. G. P. & Borysiewicz, L. K. *Immunology* **80**, 625–632 (1993).
15. Tay, C. H., Welsh, R. M. & Brutkiewicz, R. R. *J. Immunol.* **154**, 780–789 (1995).
16. Held, W., Dorfman, J. R., Wu, M. F. & Raulet, D. H. *Eur. J. Immunol.* **26**, 2286–2292 (1996).
17. Scalzo, A. A. *et al. J. Immunol.* **149**, 581–589 (1992).
18. Olsson, M. Y. *et al. Proc. Natl Acad. Sci. USA* **92**, 1649–1653 (1995).

Daedalus

Water, water everywhere

All our water comes from the air, usually by way of rain. Daedalus now wants to extract water vapour from the air directly. He points out that the vapour pressure of a solution declines in strict proportion to the concentration of dissolved molecules at its surface — a fact exploited in the determination of molecular weights. If the vapour pressure of an aqueous solution falls below that of the water in the air, water vapour will condense into the liquid (which is why you should not leave the lid off a jar of syrup or condensed milk).

These products, however, are not ideal water-traps. Solutions of surfactants should be better. Their molecules cluster preferentially at the water surface, where the concentration can be hundreds of times higher than in the bulk. So quite dilute solutions should have very low vapour pressures. Indeed, one very surface-active substance, cetyl alcohol, has been added to reservoirs, ostensibly to hinder evaporation. But Daedalus reckons it works by the converse mechanism. It doesn't stop evaporation; it encourages condensation.

Aided by this new insight, DREADCO chemists are seeking better molecules for the job. They are measuring the solution vapour pressures of powerful fluorinated surfactants and non-ionic detergents, seeking the largest possible pressure drops. Their goal is an additive for reservoirs and ponds which will concentrate at the surface strongly enough to keep vapour pressure of the pond well below that of the air. The pond will then absorb humidity all the time, even without rain. Sadly, some surfactants can sink ducks and other waterfowl by wetting their feathers. A silicone duck-grooming treatment may be needed.

Water from such a pond will, of course, contain traces of surfactant. To extract and recycle it, the team is developing a large-scale froth-blower. Air is bubbled through the detergent-contaminated water; it foams up, and its surface-active molecules concentrate in the vast surface area of froth. The froth is blown off, and collapsed by heating; the resulting strong detergent solution is returned to the reservoir.

Britain's unreliable water companies will rush to install the system. But their troubles will not be over. Why bother with piped water, if you can extract your own from a small pond or tank fitted with a DREADCO detergent system? Sprawling reservoirs and vast grids of leaky pipes will no longer be needed. Only heavy industry, with its concentrated usage, will buy piped water at all.

David Jones

Obituary

Robert Henry Dicke (1916–97)

Physicist whose work led to the discovery of the cosmic microwave background

In the exuberant world of present-day gravity physics, it is hard to picture research four decades ago when you could count the active people on your fingers. Bob Dicke, who died on 4 March, was a leader in this change, most famously in the discovery of thermal radiation from the Big Bang, which was detected by a Dicke radiometer and interpreted as part of his programme of improving the empirical basis for gravity physics.

Dicke's radiometer was a product of wartime work on radar at the MIT Radiation Laboratory. Its Dicke switch, which vastly improves sensitivity by comparing the signal with a stable reference, is a standard tool for radioastronomy. The radiometer showed that collisions broaden the absorption lines of atmospheric water vapour, explaining why the wartime drive to shorter wavelength for better radar resolution failed at one centimetre. It detected thermal radiation from the Moon, and showed that space could not be warmer than 20 K above absolute zero. Two decades later, Dicke radiometers showed that the temperature of the present-day Universe is actually 2.7 K.

After the war, Dicke continued his work on quantum optics at Princeton University. He showed that inert atoms added to a gas cell suppress first-order Doppler shifts by shortening the mean free path of radiating atoms, even though they have no effect on the thermal velocities. The physics is the same as in the Mössbauer effect: as in a crystal lattice, the confinement allows the atom only discrete values of kinetic energy, and unless the recoil is strong enough to force a quantum jump, the emitted photon gets all the decay energy.

In the mid-1950s Dicke turned to gravity physics. At the time, Einstein's general relativity was the accepted theory of gravity, but Dicke recognized that the experimental basis for this theory was weak and could be considerably improved by applying the great advances in technology of measurements that grew out of war research. Dicke's improvement of the classic Eötvös experiment showed that the gravitational accelerations of different test masses agree to three parts in 10^{11} . Another experiment, with his student Lloyd Kreuzer, used a body floating in a

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

fluid of a different composition to test the equivalence of passive and active gravitational mass (respectively the response to and the source of gravity). At neutral buoyancy, the passive mass densities are the same in the body and the fluid. The lack of a detectable change in the active mass distribution when the body is moved in the fluid shows that the two measures of mass are very nearly the same.

Dicke was fascinated by the idea that the strength of gravity might be evolving, and considered many tests for the effect. For example, a stronger gravitational interaction would cause faster nuclear burning in stars, possibly accounting for the fact that some stars appeared older than the expansion time since the Big Bang. With Carl Brans, a student of John Wheeler, Dicke proposed a generalized theory in which the strength of the gravitational interaction decreases as the Universe expands. Variants of the Brans–Dicke theory figure in models for the very early Universe, when it was tens of orders of magnitude smaller than it is now, but no longer for gravity physics at the present epoch. Current results from tests that Dicke and a few others pioneered in the 1950s, including the deflection and frequency shift of radiation near mass concentrations and the motions of planets and pulsars, are consistent with general relativity rather than the Brans–Dicke generalization.

Still deeper probes of general relativity are under way, using the condition that the distribution of the thermal radiation from the Big Bang should be consistent with that of the mass fluctuations traced by the galaxies. In the mid-1960s, Dicke led in the discovery of the key to this test, the thermal radiation. At the time, he was at most only vaguely aware of work

suggesting a hot Big Bang. Richard Tolman had shown that if the Universe expanded from a hot dense state it would be filled with thermal radiation that cools with the expansion but keeps its characteristic thermal spectrum. George Gamow added to the picture by showing that thermonuclear reactions in a hot young Universe could produce a reasonable helium abundance. Ralph Alpher and Robert Herman refined Gamow's estimate and extrapolated it to predict the present temperature, 5 K, not far from the actual value of 2.7 K. A measurement did not occur to Gamow's group. It did to Dicke, who was following another idea.

In a high-density Universe, gravity eventually stops the expansion. The Universe then collapses, and if the collapse ends in a bounce and another expansion, we have an oscillating Universe. The bounce might thermalize starlight from the last cycle. Dicke persuaded Peter Roll and one of us (D.T.W.) to build a radiometer, improved over the original by a liquid-helium reference for absolute calibration, to look for this fossil. D.T.W. gave the other of us leave to present a colloquium on the search and its significance (on the grounds that the experiment would be done before any competition could get started). Word travelled from the symposium through yet another of Dicke's former students, Ken Turner, eventually to Arno Penzias and Robert Wilson at Bell Labs, who had been seeking the cause of unexpected noise in a Dicke radiometer aimed at the sky. News of the Roll–Wilkinson experiment led them to think the noise might be extraterrestrial. The Princeton measurement, at half the wavelength of the Bell Labs result, soon gave evidence that the spectrum is thermal, as expected if this is the fossil of the Big Bang (whether or not it is thermalized starlight). And out of this grew a small industry, studying our Universe as it is now and used to be.

Bob Dicke was an enthusiastic pianist, but he loved science even more. His frequent request was, "tell me something new". In the early days he held research seminars on Friday evenings; we complained but attended because the physics was too fascinating to miss. He probably knew that we called ourselves Dicke-birds — it fitted his quiet good humour, which kept us from taking ourselves too seriously while always remembering that we had better take the physics very seriously.

P. J. E. Peebles and D. T. Wilkinson

P. J. E. Peebles and D. T. Wilkinson are in the Department of Physics, Princeton University, Princeton, New Jersey 80544, USA.

Source of volcanic glass for ancient Andean tools

Ancient hunters in southern Peru used spear points and knives made of volcanic glass (obsidian). We have located the source of the vast majority of this obsidian; a quarry on the slopes of the Cotallalli volcano in the south central Andes. Chemical analysis of obsidian from this source as well as of obsidian tools from Peru and Bolivia documents the use of Cotallalli obsidian throughout the south central Andes.

Obsidian was the preferred material for stone knives, lance and arrow points, and other tools throughout the prehistoric world because of its predictable fracturing characteristics and sharp working edge. Obsidian sources are compositionally homogeneous, so chemical analysis of obsidian artefacts can be used to infer the existence of prehistoric trade with great certainty. Obsidian tools have been found throughout Peru and Bolivia, but chemical analysis indicates that these tools derive from only a handful of sources. The location of the Quispisisa quarry that supplied obsidian to northern and central Peru has been known for years, but the source of Titicaca-basin-type obsidian, the most abundant type used in the south central Andes, has remained elusive. We now report the re-discovery of the source, which is situated on the flanks of the Cotallalli volcano overlooking the Colca Valley, in Arequipa, Peru.

The Colca Valley is located in the western cordillera of the Andes in southern Peru (Fig. 1). The western range consists of volcanic lava flows and pyroclastics that overlie deeply folded Palaeozoic and Mesozoic sedimentary levels. More than 50 volcanoes exist in the Colca River drainage, but very few volcanoes produce obsidian. The Colca Valley was carved by the Colca River, which rises close to Arequipa's eastern border with Puno¹ and is the result of glacial melting during the end of the Pleistocene epoch².

In 1991 we noted numerous obsidian tools and waste flakes, many of which were 2–4 cm long, at archaeological sites in the Colca Valley. Large unused flakes implied that costs of transport and labour to produce obsidian must have been low, further suggesting a local source. In March 1994, the quarry was found, with the help of P. Huaracha, at 4,850 m on the flanks of the extinct Cotallalli volcano. Early flint-knappers used obsidian from at least one other source in Arequipa; the Umasca quarry, a source of Cuzco-type obsidian, was located in 1990 by P. B. Trawick and I. Perez on the heights of Huaychahne between Puica and

Alca in the Cotahuasi Valley^{3,4} (E. Chavez, personal communication).

Prehistoric long-distance trade occurred between the Pacific coast and the altiplano (high plains) around Lake Titicaca. Some trade routes went through the Moquegua Valley of southern Peru⁵; at least one trade route from the coast to the altiplano passed through the Colca Valley, because Cotallalli obsidian has been recovered 200 km away on the shores of Lake Titicaca. Obsidian from the Cotallalli quarry was traded throughout the south central Andes, including the Lake Titicaca area — it was named Titicaca-basin-type obsidian because its occurrence throughout the lake basin led researchers to assume that the source was also in this basin⁶. Cotallalli obsidian was undoubtedly used for thousands of years in the Colca Valley before it was taken elsewhere — it has been recovered from Pre-Ceramic sites near Sumbay⁴ and at Tumuku⁶ (Fig. 1). The presence of Cotallalli obsidian at sites dating from the Late Pre-Ceramic (2500–1800 BC) to the Late Intermediate Period (AD 1000–1476) throughout the areas of Cuzco and the altiplano around Lake Titicaca implies that long-distance trade in obsidian between the Colca Valley and the altiplano may have existed as early as 2500 BC.

The Tiwanaku state, with its capital on the southern shores of Lake Titicaca, dominated much of Bolivia and southern Peru during the Middle Horizon (AD 600–1000).



Figure 1 Archaeological sites that used obsidian from the Cotallalli quarry between 2500 BC and AD 1532. Obsidian from the quarry was carried along trade routes extending from the Pacific coast along the Colca drainage to the Lake Titicaca basin, and from there to Cuzco. Information on archaeological sites with Cotallalli obsidian is from S. O. B. (unpublished data), M. D. G. and M. G. (unpublished data), B. Owen (personal communication), and ref. 6.

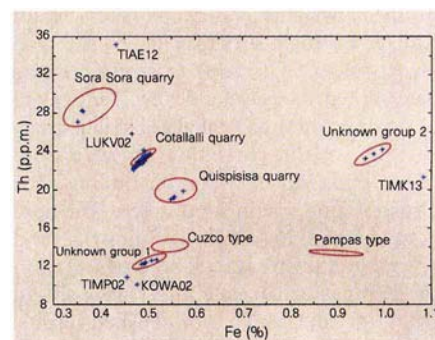


Figure 2 Bivariate plot of thorium (Th) and iron (Fe), comparing obsidian source groups (95% confidence ellipses) with individual artefacts (+ symbols). Eighty-seven artefacts from Tiwanaku sites on the shores of Lake Titicaca were analysed. Ellipses were calculated from source samples in the Missouri University Research Reactor (MURR) database. Individually labelled artefacts do not correspond to any of the known groups. The cluster for the Cotallalli quarry represents 66 of the artefacts.

Strangely, Tiwanaku ceramics and architecture have not been found in the Colca Valley, either by us or other researchers (M. Malpass and M. Neira, personal communication; ref. 7). Thus, control of the Cotallalli obsidian source during the Middle Horizon may have been held by a local group.

We have analysed 87 obsidian artefacts from Tiwanaku sites by instrumental neutron activation analysis to develop a chemical signature using standard techniques (details available on request). Obsidian artefacts were selected from the sites of Lukurmata, Konko Wankane, Chiripa, the ceremonial site of Tiwanaku and rural sites

in the Tiwanaku Valley, all sites except Chiripa having been excavated by the Wila Jawira project. Of the 87 artefacts, 66 were determined to be chemically identical to obsidian from the Cotallalli obsidian quarry (Fig. 2); the remaining 21 were from nine different sources, including the Quispisisa quarry in central Peru, the Sora Sora quarry in southwest Bolivia, and several unknown sources; no artefacts from the Pampas or Cuzco sources were found (M. D. G. and M. G., unpublished results; ref. 8). Obsidian waste flakes are most abundant in the ceremonial sectors of Tiwanaku (Akapana pyramid and Mollo Kontu mound), and in areas concentrated around the civic-ceremonial centre. Cotallalli obsidian represents at least three-quarters of the obsidian used in Tiwanaku sites.

In conclusion, obsidian from the Cotallalli quarry was carried along trade routes that extended from the Colca Valley to the altiplano and Lake Titicaca and from there to Cuzco; it was used in prehistoric sites in these areas for at least four-and-a-half thousand years. Although several minor obsidian sources in the south central Andes have yet to be located, the location of the Cotallalli quarry, the most important source in southern Peru, has now been determined.

Sarah Osgood Brooks

Department of Geography,
University of Wisconsin,
384 Science Hall,
550 North Park Street,
Madison, Wisconsin 53706-1491, USA
e-mail: sbrooks@geography.wisc.edu

Michael D. Glascock

Research Reactor Center,
University of Missouri,
Columbia, Missouri 65211, USA

Martin Giesse

Department of Anthropology,
University of Chicago,
1126 East 59th Street,
Chicago, Illinois 60637, USA

Infanticide caused by hunting of male bears

With few exceptions¹, models used to predict the effect of harvesting on animal populations disregard males². However, the killing of adult males can reduce population growth if, for example, immigrant males that replace a removed resident male kill the young^{1,3}. In a field study of brown bears, *Ursus arctos*, we have found that killing one adult male had a population effect equivalent to killing 0.5 to 1 adult females. This may be a general phenomenon in species showing this type of infanticide.

Biologists rarely consider the population consequences of a disrupted social structure when adult males die⁴. But an immigrating male replacing a dead individual may increase his reproductive success by killing existing cubs, as this can shorten the interval to a female's next conception³. We observed that young were lost primarily during the breeding season in May–June (15 of 20 cubs lost, $\chi^2_c = 15.96$, d.f. = 1, $P = 0.0001$). This loss shortens the time to next conception — 8 of 10 females that lost all young gave birth the following year, compared with none of 40 that successfully raised cubs ($\chi^2_c = 32.37$, d.f. = 1, $P < 0.0001$).

The sexually selective infanticide hypothesis³ predicts that survival of cubs less than 1 year old would be lower after a resident adult male bear is killed. We tested this using a retrospective experiment in two study areas in Sweden⁵ between 1985 and 1995. In the treatment area (south), hunters killed adult (at least 5 years old) males in 4 of the 11 years: 1989 (1 male), 1990 (1), 1991 (2), and 1992 (1). They killed none in the control area (north). We documented disappearance, assumed to mean death, among 124 unmarked cubs accompanying radio-tagged females during the period 1988–95. Bears commonly kill cubs⁶ — we observed one death by a male, and suspected another on the basis of tracks. The hunting season was in autumn, so immigration of a new male and lowered cub survival would be first observed during the next spring.

Cub survival was lower in the treatment area (0.72, $N = 74$) than the control area (0.98, $N = 50$; $\chi^2_c = 12.48$, d.f. = 1, $P = 0.0004$). Also in the treatment area, cub survival was lower both 0.5 and 1.5 years after adult males had been killed, but not after 2.5 years (Fig. 1). This indicates that the social organization of adult males was unstable for 1.5 years after losing a resident male. Cub survival was 1.00 in the treatment area and 0.98 in the control area when no adult male had been killed 1.5 years earlier, suggesting that established adult males killed few cubs.

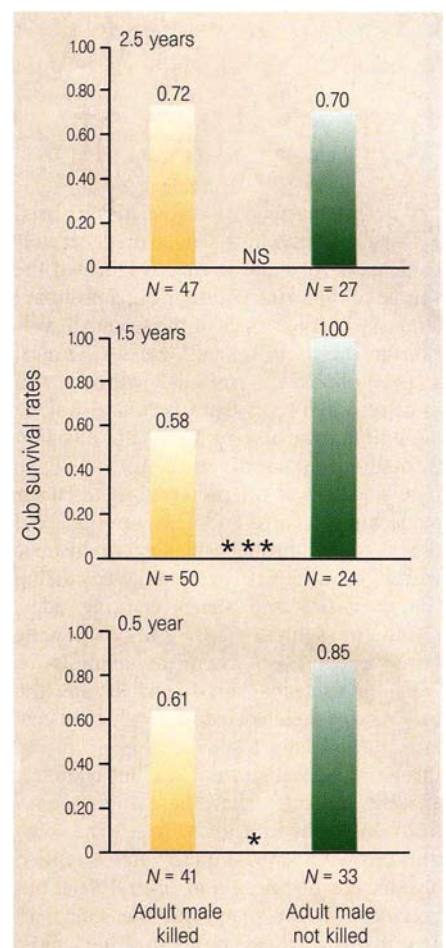


Figure 1 Comparison of brown bear cub survival rates in their year of birth in relation to whether adult males were killed (yellow bars) or not killed (green bars) in the area 0.5, 1.5 and 2.5 years before the birth of the cubs. Annual survival is indicated at the top of each bar. N = sample size; *, $P = 0.016$ ($\chi^2_c = 5.78$, d.f. = 1); ***, $P = 0.0005$ ($\chi^2_c = 12.08$, d.f. = 1); NS, not significant.

Our results cannot be explained by bear density or female condition. Cub survival was not correlated with year in this increasing bear population (logistic regression, $R = 0.000$, $\chi^2 = 0.084$, d.f. = 1, $P = 0.77$) and, unexpectedly, was negatively correlated with spring body mass of adult females ($R = -0.18$, $\chi^2 = 5.32$, d.f. = 1, $P = 0.02$).

We used computer modelling to show that the lowered cub survival reduced the population growth rate (λ) by 3.4%, from 1.18 to 1.14, and net reproductive output by 30%, in spite of a shortened litter interval. The 25–60 adult females on the 11,200 km² treatment area produce 25–60 cubs annually. A survival rate of 0.58 (Fig. 1) means that 11–25 cubs were lost, or 45–100% of the net reproductive output per female (24.6 cubs at $\lambda = 1.18$).

Bear populations are sensitive to loss of adult females^{2,7,8}. However, the effect of shooting adult males on cub survival is controversial^{8,9} and poorly documented¹⁰. We show that the population consequence of

1. Inventario, Evaluación y Uso Racional de los Recursos Naturales de la Costa: Cuenca de Río Camaná Majes (Oficina Nacional de Evaluación de Recursos Naturales, Lima, 1973).
2. Sandor, J. in *Soils in Archaeology* (ed. Holliday, V. T.) 217–245 (Smithsonian Inst., Washington DC, 1992).
3. Trawick, P. B. thesis, Yale Univ. (1994).
4. Burger, R. & Asaro, F. in *Emergencia de la Civilización en los Andes* (ed. Burger, R. L.) 189–231 (Univ. Nac. Mayor de San Marcos, Lima, 1992).
5. Stanish, C. in *Ecology, Settlement, and History in the Osmore Drainage, Peru* (eds Rice, D. S., Stanish, C. & Scarr, P.) 303–320 (Brit. Archaeol. Rep., Int. Ser. 545 ii, Oxford, 1989).
6. Burger, R. L. & Asaro, F. *Trace Element Analysis of Obsidian Artifacts from the Andes: New Perspectives on Pre-Hispanic Economic Interaction in Peru and Bolivia* (Lawrence Berkeley Laboratory, Berkeley, 1977).
7. Vera Cruz, P. *Cronología y Corología de la Cuenca del Río Camaná-Majes-Colca-Arequipa* (Tesis Licenciado en Arqueología, Univ. Católica Santa María, Arequipa, 1989).
8. Giesse, M. in *Tiwanaku and Its Hinterland: Archaeological and Paleocological Investigations in the Lake Titicaca Basin of Bolivia* Vol. 2 (ed. Kolata, A. L.) (Smithsonian Inst., Washington DC, in the press).

harvesting one adult male brown bear corresponds to that of harvesting 0.5–1.0 adult females, although the per capita effect would decline with increasing killing of males. This may be a general phenomenon and should be considered when managing threatened or endangered populations of other species that may exhibit sexually selected infanticide.

Jon E. Swenson

Norwegian Institute for Nature Research,
Tungasletta 2, N-7005 Trondheim, Norway
e-mail: jon.swenson@nina.nina.no

Finn Sandegren, Arne Söderberg

Research Division, Swedish Hunters' Association,
Box 7002, S-750 07 Uppsala, Sweden

Anders Bjärvall, Robert Franzén

Environment Protection Agency,
S-106 48 Stockholm, Sweden

Petter Wabakken

Environmental Protection Division,
County Governor's Office, Fylkeshuset,
N-2300 Hamar, Norway

- Starfield, A., Furniss, P. R. & Smuts, G. L. in *Dynamics of Large Mammal Populations* (eds Fowler, C. W. & Smith, T. D.) 121–134 (Wiley, New York, 1981).
- Miller, S. D. *Int. Conf. Bear Res. Management* 8, 357–373 (1990).
- Hrdy, S. B. *Ethol. Sociobiol.* 1, 13–40 (1979).
- Lott, D. F. *Intraspecific Variation in the Social Systems of Wild Vertebrates* (Cambridge Univ. Press, 1991).
- Swenson, J. E. *et al. Biol. Conserv.* 70, 9–17 (1994).
- Taylor, M. (ed.) *Int. Conf. Bear Res. Management* 8, Monogr. Ser. 3 (1994).
- Knight, R. R. & Eberhardt, L. L. *J. Wildl. Management* 48, 1434–1438 (1984).
- McCullough, D. R. in *Dynamics of Large Mammal Populations* (eds Fowler, C. W. & Smith, T. D.) 173–196 (Wiley, New York, 1981).
- Wielgus, R. B. & Bunnell, F. L. *Biol. Conserv.* 67, 161–166 (1994).
- Miller, S. D. *Wildl. Soc. Bull.* 18, 462–467 (1994).

A plant cyclin-dependent kinase inhibitor gene

Mammalian cyclin-dependent kinase (CDK) inhibitors are important in the regulation of the cell cycle and in development¹. Using the yeast two-hybrid system² we have identified a CDK inhibitor gene in *Arabidopsis thaliana*, which we have named *ICK1*. The gene product has unique structural features and kinase inhibitory properties. Although the carboxy-terminal domain of *ICK1* shares limited similarity with the mammalian *p27^{Kip1}* kinase inhibitor^{3,4} the remainder of the deduced *ICK1* sequence shows no similarity to known proteins. Recombinant *ICK1* specifically inhibits the histone H1 kinase activity of plant Cdc2-like kinases, but has no effect on human *p34^{cdc2}* activity.

Little is known about the molecular mechanisms underlying plant cell-cycle-related processes, although two *cdc2* homologues^{5,6} have been identified from *A. thaliana*. The functional importance of *cdc2a* has been demonstrated by its ability

to complement partially the yeast *cdc2* mutants^{3,6}, as seen from the correlation of its expression with the 'competence' of the cell to divide⁷ and by the observation that expression of a dominant negative *cdc2a* mutant resulted in cell-cycle arrest⁸. To identify genes encoding proteins that interact with the Cdc2a protein, a yeast two-hybrid 'bait' construct harbouring *A. thaliana cdc2a* complementary DNA was used to screen a library. Analysis of 68 positive clones identified three classes of proteins that interact with Cdc2 kinase (hence the name *ICK*), 55 of these clones represented various lengths of *ICK1*.

The longest *ICK1* open reading frame predicts a polypeptide of 191 amino acids (relative molecular mass 22,200). *ICK1* was similar across a 31-amino-acid region to *p27^{Kip1}*, a recently described mammalian CDK inhibitor^{3,4}. *ICK1* also shares similarity in this region, to a lesser degree, with *p21^{Cip1}*, another mammalian CDK inhibitor⁹. *ICK1* is, however, structurally distinct from mammalian inhibitors in several respects. First, the remainder of the deduced *ICK1* sequence shows no significant similarity to mammalian CDK inhibitors or other proteins. Second, the consensus sequence is located in the carboxy-terminal region of *ICK1* in contrast to its amino-terminal location in both *p27^{Kip1}* and *p21^{Cip1}* (Fig. 1). Third, *ICK1* is predicted to contain a strong coiled-coil domain in the central region of the protein (Fig. 1). Fourth, *ICK1* lacks a typical bipartite nuclear localization sequence found in both *p27^{Kip1}* and *p21^{Cip1}*.

To analyse the function of *ICK1* as a kinase inhibitor, recombinant *ICK1* was expressed using a modified pRSETB expression vector (Invitrogen) and purified as a (His)₆ fusion protein before use in the *in vitro* kinase assays. Recombinant *ICK1* at nanomolar concentrations was found to inhibit the histone H1 kinase activity of the partially purified *A. thaliana* Cdc2-like kinases (Fig. 2a). The inhibition reached a maximum plateau level at about 80% of the total activity. The incomplete inhibition was either due to the inability of *ICK1* to inhibit the activities of all the CDK complexes purified, or perhaps to other unidentified kinases present in the preparation that are unaffected by *ICK1*.

Comparable concentrations of recombinant *ICK1* had no detectable effect on the



Figure 1 Structural comparison between *ICK1* (GenBank accession number U94772) and *p27^{Kip1}*. Regions of similarity between the two proteins are indicated by black shading and the predicted strong coiled-coil structure in *ICK1* by grey shading. The coiled-coil structure was predicted using MTK matrix (ISREC server <http://ulrec3.unil.ch/>, COILS version 2.1).

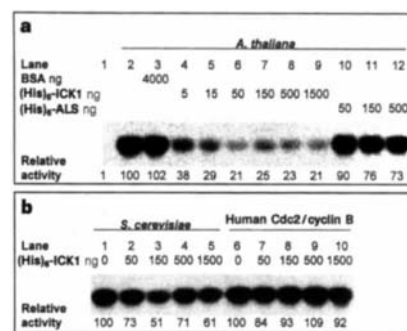


Figure 2 a, Effect of (His)₆-ICK1 on the histone H1 kinase activity of (lanes 4–9) *A. thaliana* p13^{suc1}-associated kinases, compared with (lane 3) bovine serum albumin and (lanes 10–12) recombinant (His)₆-acetolactate synthase (His)₆-ALS). Lane 1, control with an equal amount of beads lacking plant extract. b, Effect of (His)₆-ICK1 on the histone H1 kinase activity of (lanes 1–5) *S. cerevisiae* p13^{suc1}-associated kinases and (lanes 6–10) human Cdc2/cyclin B complex. p13^{suc1}-affinity chromatography was used to purify p13^{suc1}-associated kinases from *A. thaliana* (150 µg protein per treatment) and *S. cerevisiae* (80 µg protein per treatment). Human Cdc2/cyclin B complex was purified from baculovirus-infected insect cells. Recombinant *ICK1* or a control protein was added and incubated (tumbling slowly, 1.5 h, 4 °C). The kinase reaction was then initiated (1 µg µl⁻¹ histone H1 (Sigma), 25 µM ATP, 0.05 µCi µl⁻¹ ³²P-ATP, 20 min). Denatured supernatant was resolved by SDS-PAGE. ³²P-labelled histone H1 was quantified (Molecular Dynamics PhosphorImager) and relative activities are shown beneath each lane.

histone H1 kinase activity of purified human Cdc2/cyclin B complex, and showed only a minor effect on that of partially purified *S. cerevisiae* CDC28 kinase (Fig. 2b). As it is known that *p27^{Kip1}* inhibits mammalian Cdc2 activity, the finding that relatively high concentrations of *ICK1* failed to inhibit human Cdc2/cyclin B kinase activity may point to a functional importance for the structural divergence of *ICK1*.

Although studies of maize endosperm development have indicated that CDK inhibitor activity might exist¹⁰, *ICK1* represents the first CDK inhibitor gene to be identified from a plant. We believe that the functional aspects of our approach were important in identifying *ICK1*, as it shows a high degree of sequence divergence from its counterparts in other species. The identification of *ICK1* provides a tool for study of the roles of plant CDK inhibitors in cell cycle, growth and development.

Hong Wang, Larry C. Fowke

Department of Biology, University of Saskatchewan,
Saskatoon S7N 5E2, Canada
e-mail: hwang@pbi.nrc.ca

William L. Crosby

NRCC Plant Biotechnology Institute,
Saskatoon S7N 0W9, Canada

- Harper, J. W. & Elledge, S. J. *Curr. Opin. Genet. Dev.* **6**, 56–64 (1996).
- Fields, S. & Song, O.-K. *Nature* **340**, 245–246 (1989).
- Polyak, K. *et al. Cell* **78**, 59–66 (1994).
- Toyoshima, H. & Hunter, T. *Cell* **78**, 67–74 (1994).
- Ferreira, P. C. G., Hemerly, A. S., Villarreal, R., Van Montagu, M. & Inzé, D. *Plant Cell* **3**, 531–540 (1991).
- Hirahama, T., Imajuku, Y., Anai, T., Matsui, M. & Oka, A. *Gene* **105**, 159–165 (1991).
- Martinez, M. C., Jørgensen, J.-E., Lawton, M. A., Lamb, C. J. & Doerner, P. W. *Proc. Natl Acad. Sci. USA* **89**, 7360–7364 (1992).
- Hemerly, A. *et al. EMBO J.* **14**, 3925–3936 (1995).
- Harper, J. W., Adami, G. R., Wei, N., Keyomarsi, K. & Elledge, S. J. *Cell* **75**, 805–816 (1993).
- Grafi, G. & Larkins, B. A. *Science* **269**, 1262–1264 (1995).

Determining cortical landscapes

The folded surface of the cerebral cortex contains a mosaic of functionally distinct areas, connected by a network of nerve fibres in the sub-cortical white matter. This network determines, in part, the computations that the cortex performs. The folds help to fit the large sheet-like cortex into a compact space¹, but this does not explain the species specificity of their pattern, or their characteristic relationship with the locations of the cortical areas². Here I present data in support of

Van Essen's Hypothesis² to explain these features.

Local wiring — preferential connectivity between nearby areas of the cortical sheet — is a simple strategy that helps keep cortical connections short^{3–5}. In principle, efficient cortical folding could further reduce connection length, in turn reducing white matter volume and conduction delays^{2,6}. Such an arrangement should match convex folds (gyri) with groups of heavily connected areas, and concavities (sulci) with sparsely connected regions. Van Essen² has suggested that competing mechanical tension in the axons of the developing cortical network produces such efficient folding.

Cats⁵ and macaques⁷ have roughly 70 cortical areas in each folded hemisphere, linked by around 1,000 connections. The complexity makes it possible to find subsets of connections to support virtually any theory of cortical organization. Therefore, I analysed collated anatomical data quantitatively to see if the 3-dimensional folding of the cortex is in fact related to wiring. I examined the folding and connections between areas that are 'nearest neighbours' or 'next-door-but-one'^{3–5} on the sheet. In both animals, connectivity was significantly denser between nearby areas on the same gyrus than between nearby areas sep-

arated by sulci (Fig. 1, Table 1).

In the cat, where data are available on the relative strengths of connections, neighbouring and next-door-but-one areas within gyri had more moderate and strong connections whereas those separated by sulci had more sparse connections ($P < 0.05$ by binomial tests). Thus the development of the cortex does seem to coordinate folding with connectivity in a way that produces smaller and faster brains. Many mechanisms could conceivably account for this (for example, growth factors from extrinsic connections), but the result is quantitatively consistent with Van Essen's theory of morphogenesis².

Despite the presence of local wiring^{3,5} and efficient folding^{2,6}, 53% of connections in the cat link areas that are neither neighbours nor next-door-but-one, whereas 42% of such areas remain unconnected (Table 1). Furthermore, 42% of the connections between nearest neighbours and next-door-but-one areas cross sulci, whereas 32% of such areas within gyri are unconnected. Similar figures apply to the macaque (Table 1). Thus natural selection has produced a 3-dimensional arrangement of areas that is not optimal for minimizing wiring. Indeed, the connective architecture could be perfectly accommodated only if local wiring and folding were able to take place in more than three dimensions^{5,7,8}. Brain design represents a compromise between efficient wiring and other factors⁸.

Imagine a cortex in which it were possible to shuffle the areas and folds while keeping connections the same. Shuffling the areas (abolishing local wiring) would compromise compact wiring more than shuffling the folds (abolishing efficient folding; Table 1). Morphologically spectacular folding may, thus, 'tune up' a cortex that is already reasonably compact from local wiring, particularly in larger brains where wiring 'costs' are greater^{1,9}. I anticipate, however, that local wiring rules apply even in those cortices that do not fold.

Jack W. Scannell

Neural Systems Group,
Psychology Department,
University of Newcastle,
Newcastle-upon-Tyne NE1 7RU, UK
e-mail: j.w.scannell@ncl.ac.uk

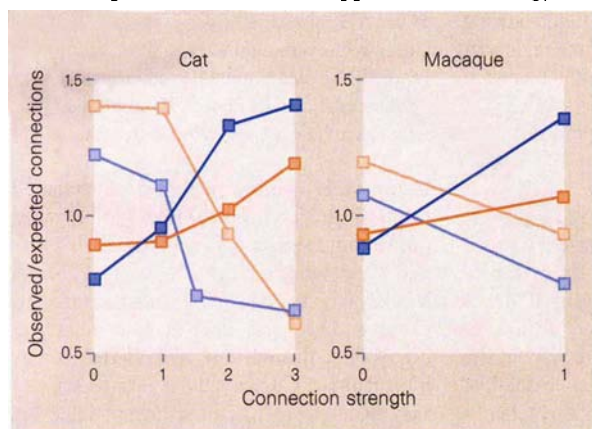


Figure 1 Local connections are denser within gyri (dark blue and orange) than across sulci (pale blue and orange). Nearest neighbours plus next-door-but-one areas (blue) and neighbouring areas alone (orange) are compared. The ratio of observed/expected connections is plotted against connection weight. The differences between observed and expected connections are significant (see Table 1).

Table 1 Cortical connections, observed and expected

	Weight	(1) NN and NDB1		(2) NN		(3) All	
		Gyri	Sulci	Gyri	Sulci	NN and NDB1	Non-local
Cat	0	105/138	172/139	31/37	23/17	277/452	1,515/1,338
	1	66/63	70/63	16/19	12/9	126/99	266/293
	2	122/94	66/94	71/67	27/31	188/85	147/250
	3	47/35	24/36	34/29	8/13	71/23	23/70
		$P < 0.0001$		$P < 0.04$		$P < 0.0001$	
Macaque	0	211/255	344/300	70/78	51/43	555/751	3,498/3,301
	1	195/151	132/176	122/114	53/61	327/130	378/574
		$P < 0.0001$		$P < 0.04$		$P < 0.0001$	

The cortex shows efficient folding and local wiring. Number of observed (bold) over number of expected connections are compared. Comparisons are (1) within gyri compared with across sulci connections between nearest neighbour (NN) and next door but one (NDB1) areas, (2) within gyri compared with across sulci connections between NN areas, (3) all NN and NDB1 compared with connections between all non-local areas^{2–5}. Weight gives the ordinal rank of the strength of connection. Cat⁵ connection density ran from '0' (no connection) to '3' (strong connection). Macaque⁷ connections were either '0' (absent) or '1' (present). Probabilities were calculated using χ^2 . If local wiring and efficient folding were absent, the observed values in the table would equal the expected values, hence the table shows that, for NN and NDB1 areas, local wiring contributes more than folding to efficient cortical connectivity. Full details of anatomical data are available at <http://www.psychology.ncl.ac.uk/jack/gyri.html>.

- Welker, W. in *Cerebral Cortex* Vol. 8B (eds Jones, E. G. & Peters, A.) 3–136 (Plenum, New York, 1990).
- Van Essen, D. C. *Nature* **385**, 313–318 (1997).
- Young, M. P. *Nature* **358**, 152–155 (1992).
- Cherniak, C. J. *Neurosci.* **14**, 2418–2427 (1994).
- Scannell, J. W., Young, M. P. & Blakemore, C. J. *Neurosci.* **15**, 1463–1483 (1995).
- Griffin, L. J. *Theoret. Biol.* **166**, 261–273 (1994).
- Young, M. P. *Proc. R. Soc. Lond. B* **252**, 13–18 (1993).
- Young, M. P. & Scannell, J. W. *Trends Neurosci.* **19**, 413–415 (1996).
- Ringo, J. L. *Brain Behav. Evol.* **38**, 1–16 (1991).

The policy intellectual

David E. Lilienthal: The Journey of an American Liberal

by Steven M. Neuse

University of Tennessee Press: 1996.

Pp. 406. \$40

Howard P. Segal

Steven Neuse has rescued David Lilienthal from an undeserved growing obscurity. When Lilienthal died aged 81 in 1981, *The New York Times* ran his obituary on its front page and praised him on its editorial page as "a special citizen" who epitomized public service and administrative competence. Lilienthal had been one of the three original directors of the Tennessee Valley Authority (TVA), established in 1933 by Franklin Roosevelt early in his presidency. After several years of enormous internal and external controversy over its mission and power, TVA eventually became the nation's largest electrical utility and the most popular New Deal agency. Lilienthal remained a director until 1946, spending the last five years as board chairman.

During that time, and later as the first chairman of the Atomic Energy Commission, Lilienthal exemplified what I would call the policy intellectual. Unlike the so-called public intellectual, the policy intellectual seeks involvement in rather than distance from decision-making, yet is detached and perceptive enough to write about policy formulation in more than superficial ways. Admittedly, Lilienthal neither was nor claimed to be profound. He often chastised as too abstract and unreadable the works of those who did claim profundity, particularly academics. Trained as a lawyer at Harvard, Lilienthal had little patience with theory separated from practice. Nevertheless, his many speeches, articles and books — including seven volumes of journals — greatly illuminate public affairs and policy formulation in the twentieth-century United States. C. P. Snow is a roughly comparable figure for the same period in Britain; Lilienthal would have made a perfect character for a Snow novel.

Lilienthal's remarkable drive, ambition, and political and rhetorical skills accounted for much of TVA's achievement, though others also deserving credit were often overlooked. The most notable of these was Arthur Morgan, the original board chairman, fired by Roosevelt in 1938 after failing to document public charges of malfeasance against his fellow directors. Despite poor political judgement and irritating self-righteousness, Morgan, who was a flood-control engineer, was largely responsible for TVA's initial dams and irrigation systems, the agency's most visible triumphs. Lilienthal omitted Morgan's name from his *TVA: Democracy on the March* (1944), an eloquent, impassioned and influ-

ential defence of the agency as almost utopian in both its operations and its alleged 'grass-roots' democratic decision-making.

Lilienthal left the TVA he loved only because President Harry Truman persuaded him that the new Atomic Energy Commission required his leadership. From 1946 until 1950 Lilienthal tried to establish policy for the peaceful uses of atomic power amid growing Cold War tensions. This elevated his public profile still higher, but left no legacy equivalent to that of TVA. Most Americans agreed with Truman and Lilienthal that civilians rather than the military should control the atom. But most rejected as naive and impractical Lilienthal's and others' proposed international cooperation under an Atomic Development Authority, and the outright sharing of atomic secrets and safeguards with the Soviets. Still, Lilienthal proved farsighted in his opposition to the atomic arms race that quickly followed.

Drained of both energy and finances after nearly two decades of exhausting and modestly paid service, Lilienthal spent the rest of his life in the private sector. In 1955, he established the Development and Resources Corporation with several former TVA associates. By now heralded around the world as 'Mr TVA', Lilienthal worked in more than 25 countries trying to duplicate the combination of large energy and irrigation projects, technical expertise and supposed grass-roots decision-making that had made TVA so special and so successful.

Yet the undemocratic govern-

Lilienthal: epitomized public service and administrative excellence.

ments for whom the corporation primarily worked often proved so unstable and unreliable that they were poor investments. The company went bankrupt in 1979 when the Shah of Iran was deposed and the new revolutionary government refused to reimburse the corporation for the projects that, by then, accounted for nearly all its profits. Lilienthal's repeated failure to pay more than lip service to grass-roots democracy in countries whose ruling élites neither represented nor cared about public opinion came back to haunt him in Iran and elsewhere.

Neuse is a political scientist who has provided a quite readable biography based on enormous research. He is sympathetic toward Lilienthal throughout, but concedes his subject's failings as a sometimes ruthless administrator, unceasing sycophant and lifelong true believer in the power of American technology to improve the culture and politics — not just the infrastructure — of every other country. Especially moving is Neuse's

description of the way the ageing Lilienthal was willingly manipulated, by President Lyndon Johnson, to defend the Vietnam War as the precursor to TVA-like megaprojects once the United States won.

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

It estranged Lilienthal from his anti-war son for years thereafter, and added the father's name to the long list of liberal élitists and technocrats whose reputations never recovered from that débâcle.

By the time Lilienthal died, Americans' faith in unadulterated technological progress had faded and been replaced by scepticism and even pessimism about technology as panacea. In that respect, Lilienthal's increasing obscurity is not surprising. But Neuse describes a much more hopeful era with which to compare our own, and that is why his book is important and deserves a wide readership. □

Howard P. Segal is in the Department of History, University of Maine, Orono, Maine 04469-5774, USA.

Darwin on the brain

Evolutionary Psychiatry: A New Beginning

by Anthony Stevens and John Price
Routledge: 1996. Pp. 267. £47.50, \$69.95 (hbk); £14.99, \$18.95 (pbk)

Steven Rose

Alas, poor Darwin! More idle speculation and dogmatic assertion have been published in your name in the past two decades than in the full preceding century, and still the torrent continues. Evolutionary biology spawns evolutionary psychology, sociology, medicine, economics, and now psychiatry.

That the human condition is the product of evolutionary processes involving natural selection is a truism to all biologists. That our diseases and disorders take the form they do as a result of these processes, embedded as they are within human history, culture and society, would scarcely seem a startling generalization, despite the wonder that such statements seem still to generate. But, as ever, the devil lies in the details, and in the evidential basis that could turn such broad claims into useful insights.

Anthony Stevens and John Price, respectively a Jungian analyst and a psychiatrist, begin their book with the grandest of claims for the revolution in understanding that they hope to provide — but they soon stumble. In fact the book turns out to be far more an attempt to vindicate Jung's theory of archetypes by linking it to Darwin on the one hand and John Bowlby on the other than one might have anticipated from the promissory note of its preface. It is written very much in the style of a teaching text, its pages spattered with bold type somewhat randomly indicating terms to be found in the glossary.

It whisks quickly through a classical introduction to the history of (non-biological) psychiatry towards a section bravely entitled "Human Nature: its Evolution and Development", which turns out to mean an uneasy conflation of Hal Waddington, Edward

Wilson and Bowlby. Unfortunately the authors' genetics and neurobiology get a bit shaky at this point. The human genome is defined as "our innate propensities," while Paul Maclean's outdated theory of the triune brain gets yet another airing, so that later in the book the authors can make much play with our "reptilian brain" whose "emanations" thoroughly condition "our drives, inner subjective feelings, fantasies and thoughts".

Stevens and Price begin with a refreshing attempt to transcend the dreary listing of psychiatric disorders provided by the *Diagnostic and Statistical Manual* (DSM) with its enthusiasm for the minute nosological classification best regarded as itself a sort of intellectual disorder (perhaps syndromitis?). Instead they argue — surely correctly — that psychiatric distress should be considered case by case rather than given Procrustean labels. They find it hard to maintain this position, though, and for a good portion of the latter part of the book even find themselves labelling chapter subheads according to DSM criteria — "Avoidant Personality Disorder" is "DSM301.82", for instance.

They then subject each such 'disorder' to a sociobiological analysis. This is based on the view that constraints of rank and spacing, arising during humanity's presumed

healthy Environment of Evolutionary Adaptedness (or EEA — this and other acronyms litter the text), generate the dysfunctions and psychic distress that permeate modern society. In this view, our ancestors may not have been Noble, but were at least Psychically Healthy Savages.

It is here that the authors' evolutionary and anthropological fantasies are given full rein. They argue that agoraphobia is commoner in women than in men because in the EEA men hunted while women stayed at home, so women became afraid to stray long distances without their men. Obsessive-compulsive disorder "probably arose in relation to the defence of resources... security arrangements had to be frequently and thoroughly checked... with wide genetic variation, and will still be under active selection". Homosexuality may be a kin-selection adaptation by which childless gays invest more in the care of their relatives' offspring. Recognizing that this suggestion, originally made by Wilson, has no current empirical base, the authors maintain nonetheless that it would have been a necessary condition of survival "in the ancestral environment". And even today "wealthy homosexuals sometimes shock their friends by leaving the bulk of their estate... to some nephew or niece whom they may not have seen for years".

New in paperback

Pythagoras' Trousers: God, Physics, and the Gender Wars

by Margaret Wertheim

Fourth Estate, £9.99

"Wertheim takes the reader through the development of physics, starting with the mathematics and astronomy of ancient Greece, and moving through Bacon, Copernicus, Galileo and all the rest of the seventeenth-century crowd, up to the present day. As a well-researched, historical account I found it fascinating, and I was intrigued by her theme that today's male domination of physics has its roots in the historically close connections between physics and Christianity." Keith Devlin, *Nature* 379, 128 (1996).

The Oxford Book of Nature Writing

by Richard Mabey

Oxford University Press, £7.99

"Starting with Aesop and Aristotle and ending with a piece written just a decade ago by Primo Levi, there is something here for everyone." David E. Allen, *Nature* 374, 319 (1995).

Dynamic Patterns: The Self-Organization of Brain and Behaviour

by J. A. Scott Kelso

MIT Press, \$22.50, £18.95

"This charming and thought-provoking book looks at the brain and behaviour as pattern-forming dynamical systems. It captures, and

gives substance to, a new theme in neuroscience, in which complexity is in the ascendancy and good old-fashioned chaos seems a thing of the past." Karl J. Friston, *Nature* 375, 643 (1995).

Biochemical Oscillations and Cellular Rhythms: the Molecular Bases of Periodic and Chaotic Behaviour

by Albert Goldbeter

Cambridge University Press, £29.95, \$44.95

"Although it will demand some hard thinking, the book can serve as a reliable and rewarding introduction to mathematical modelling of dynamic processes in living cells for any molecular biologists beginning to see the usefulness of this tool." John J. Tyson, *Nature* 380, 213 (1996).

Rebel With a Cause: an Autobiography

by Hans Eysenck

Transaction Publishers, \$21.95, £14.95

This is a revised and expanded edition. "It must be said that autobiography is not Eysenck's natural *métier*. Perhaps because it is the least intellectually pugilistic, least controversial and tendentious, and most narrative of his works, it is less than gripping stuff. As a matter of record, it would have been a pity had it been left unwritten. But as an idiographic account of one man's life, which it explicitly sets out to be, it is disappointing." Steve Blinkhorn, *Nature* 344, 889 (1990).

This trivial combination of gossip about current social mores and wild guesses about life in the Palaeolithic does neither authors nor publisher much credit. At a period when both evolutionary theory and psychiatry are confronted by entrenched cultural opponents, they deserve better than this from their friends. A more sophisticated biosocial approach to psychic distress, treading the delicate but necessary line between biological and social determinisms, would be really welcome.

Until that is achieved, the time may have come for a moratorium on books with the word evolution in their title. □

Steven Rose is in the Brain and Behaviour Research Group, The Open University, Walton Hall, Milton Keynes MK7 6AA, UK.

Back to the roots of botany in China

A Botanical Pioneer in South West China

by Heinrich Handel-Mazzetti

(translated by David Winstanley)

David Winstanley: 1996. Pp. 192. Available from Alpine Garden Society Publications, AGS Centre, Avon Bank, Pershore, Worcestershire WR10 3JP, UK. £14 plus p&p (UK £3.15, overseas £5.50)

Mark F. Watson

The last decade of the nineteenth century and the 30 years that followed were a golden age of plant hunting in southwest China. Missionaries and government officials were the first Europeans to explore the Chinese interior and those with botanical interests soon sent excited messages back to Europe recounting the wealth of plants to be found.

It was not long before nurserymen and seed merchants despatched collectors to comb the area for novel plants for gardens. So bounteous were the gatherings of these men — notably George Forrest, Joseph Rock and Frank Kingdon Ward — that the area was dubbed the 'Mother of Gardens'.

The Austrian botanist and mountaineer Heinrich Handel-Mazzetti went to China in 1914 but, soon after he penetrated deep into northwest Yunnan, hostilities broke out in Europe and he was stranded for the duration of the First World War. He seized upon this opportunity, and despite great difficulties travelled widely in southwest China. As a scientist he was interested only in herbarium collections for systematic study and documenting their ecology, and unlike other contemporary collectors he did not introduce any new plants into cultivation.

So Handel-Mazzetti did not get entangled in the rivalry between the horticultural collectors, who were jostling for choice finds and the prestige that their introduction would bring. But he respected their work and



The Salween Valley in China, from an oil painting by Eduard Handel-Mazzetti.

sought their advice whenever he could, and in areas well covered by them he concentrated on mosses, liverworts, grasses, sedges and other overlooked groups.

Handel-Mazzetti's herbarium specimens are outstanding, not only for the quality of the plant material but also for the precision of the label data. His breadth of knowledge was extraordinary, and he is perhaps most greatly respected today as a pioneer by cryptogamists.

In 1914 Handel-Mazzetti met Forrest in Lijiang. Forrest was able to brief Handel-Mazzetti on the terrain of his planned route, suggesting areas that needed exploring and giving advice on dealing with the local people. These were not easy times for foreigners in deepest China. Not only were there hardships of travel, poor accommodation and terrible food, but there were also very real threats of hostility. The French missionary and botanist Jean-Théodore Monbeig had been murdered a few weeks before Handel-Mazzetti's arrival, and later Forrest had a narrow escape when he had to flee from Tibetan warrior-monks.

After Handel-Mazzetti returned to Vienna, Forrest suggested that he contact the Royal Botanic Garden in Edinburgh to offer exchange of herbarium specimens. Today we have a good number of Handel-Mazzetti's collections at Edinburgh. But, judging by our archive records of the frosty initial responses he received from staff, it is a wonder we have anything at all! (In 1921–22 post-war feelings against Germans and Austrians were still strong.)

Until now, Handel-Mazzetti's accounts of his remarkable travels have remained locked away in the rare German volume

Naturbilder aus Südwest-China. This book, and his later taxonomic work *Symbolae Sinicae*, are treasure chests holding critical systematic notes and historical documentation. But they are written in a very formal high German inaccessible to most botanists. David Winstanley must be congratulated on his wonderful translation which accurately conveys both the content of the original work and the sentiment of the author.

Winstanley has supplemented his translation with a fascinating biographical memoir, compiled from primary and secondary sources, which gives a fair insight into the character of the man. Also included are a very useful glossary of Chinese place names, and biographical notes on the contemporary explorers Camillo Schneider, Edward Amundsen and Roy Andrews.

Handel-Mazzetti illustrated his book lavishly with some 150 photographs (colour and monochrome). It is a shame that only a third of these have been reprinted in the present volume because of financial limitations. But the quality of the reproductions is superb as they have been printed from the original negatives, and are much better than in the original work. Handel-Mazzetti may not have been a great photographer, but his pictures are a valuable record of the vegetation before the ravages of the Maoist Great Leap Forward and subsequent developments in China. Peasant folk and rural village scenes are much the same today, and the monks in their monasteries seem trapped in time. So it is a great pity that a publishing house did not take the opportunity to fund the production of a complete reprinting of the original photographs.

Financial constraints must also explain the rather poor and very condensed quality of the print which detracts from an otherwise excellent publication.

In 1927 Augustine Henry reviewed *Naturbilder aus Südwest-China* in *Nature* (119, 667–668; 1927) in glowing terms, commenting that the book has “a strong appeal to horticulturists; and we hope that a translation into English will soon be published”. It may have taken 70 years to appear but at last we have it, and undoubtedly it will become a main reference work for botanists working on Handel-Mazzetti’s collections, as well as those interested in Chinese plants. □

Mark F. Watson is at the Royal Botanic Garden, Inverleith Row, Edinburgh EH3 5LR, UK.

A revolution in evolution

Patterns in Evolution: The New Molecular View

by Roger Lewin

W. H. Freeman: 1997. Pp. 245. £19.95, \$32.95

Svante Pääbo

It is often said that biology is becoming increasingly specialized. What this book shows is that, at least in evolutionary biology, the opposite trend is at work. Many of the questions in evolutionary biology were formulated by generations of scholars specializing in the study of particular groups of organisms. Yet molecular evolutionists apply the same techniques and interpretative tools regardless of whether they study bacteria, fungi, birds or humans. Therefore those who use molecular techniques can both talk to, and learn from, each other. As molecular data accu-

mulate, this trend will continue. Evolutionary biologists can rejoice in living in a time of synthesis, rather than fragmentation, of their field.

Over the past ten years, molecular genetic approaches have revolutionized evolutionary biology. Roger Lewin’s overview of this development begins by describing some major issues in the field, starting with the ‘molecules versus morphology’ controversy, the question of whether morphological characters or gene sequences provide a better guide to evolutionary relationships. A balanced consensus seems to be in sight: where morphological studies have yielded unequivocal answers, molecular data tend to confirm them; but where morphology has not provided a clear view, molecular studies are invaluable because they yield large numbers of clearly delineated characters with modes of change that can be modelled mathematically. Many ‘trichotomies’ or ‘explosive radiations’ may yet be resolved by molecular studies, particularly given recent technical advances that allow huge amounts of DNA sequences to be determined quickly and inexpensively.

Lewin then discusses molecular phylogenies, the ‘selection versus neutrality’ debate, and ‘molecular clocks’. Some oversimplifications here may annoy the insider, but the main message is clear: the neutral model should serve as the null hypothesis against which molecular data are tested, as it makes defined, and therefore falsifiable, predictions. Lewin goes on to describe the molecular contributions to ecology and anthropology, and the retrieval of ancient DNA.

This is an excellent and beautifully illustrated book, and the text is written in a lucid and captivating fashion. I can recommend it to anyone who wants an accessible introduction to the field of molecular evolution. □

Svante Pääbo is at the Institute of Zoology, University of Munich, D-80021 Munich, Germany.

Mind over matter

The Large, the Small and the Human Mind

by Roger Penrose

Cambridge University Press: 1997. Pp. 185. £14.95, \$19.95

Philip W. Anderson

Roger Penrose is devoted to persuading us that the three problems in science that seem to him to be the deepest and most intractable are, somehow, to be solved by the same somewhat vague inspiration. These problems are the quantum theory of gravity, measurement and decoherence in the quantum theory, and the nature of consciousness. How he manages to link these three, and in the last of them to implicate microtubules (which are objects of known structure and function that control the internal working of all of our cells, including neurons), is not at all clear.

In this respect, and many others, *The Large, the Small and the Human Mind* is similar to his previous book, *Shadows of the Mind*, of which it seems to be a digest. It embodies all of the same arguments and a selection of the same words and illustrations. *Shadows of the Mind*, in turn, elaborated at some length on some ideas contained in his previous very popular book, *The Emperor’s New Mind*.

New to *The Large, the Small and the Human Mind* are discussions from one mathematical physicist, Penrose’s former student Stephen Hawking, and two philosophers, Nancy Cartwright and Abner Shimony. Hawking explains Penrose’s ideas with admirable clarity, and equally clearly lays out the criticisms of these ideas that an empirically minded physicist is likely to have. Cartwright asks, very properly, “why not a biologist?”, by which I presume she means that biology has its own somewhat independent intellectual content that Penrose’s strong reductionist point of view ignores, and therefore it may be inadequate to deal with problems such as consciousness. I read her as suggesting that a biologist’s commentary would have been useful, and personally I agree with both points. Shimony, on the other hand, introduces a concept he calls “Whiteheadism”, which to me suggests primitive ideas such as the *élan vital* or even animism, which are foreign to the point of view of modern neuroscience.

It is clear that I am saying, with some regret, that I cannot recommend the present book to those who have read either of the previous ones, and that, whatever one may make of Penrose’s thesis, the earlier books are superior in exposition and content.

Philip W. Anderson is in the Department of Physics, Princeton University, Princeton, New Jersey 08544, USA.



Easter eggs

Before Easter slips away, here are some eggs — from the dinosaur *Spheroolithus irenensis*. For more eggs, shell out on *Fossil and Recent*

Eggshell in Amniotic Vertebrates: Fine Structure, Comparative Morphology and Classification (Special Papers in Palaeontology no. 56) by K. E. Mikhailov (The Palaeontological Association, £35).

Maximal sustained energy budgets in humans and animals

Kimberly A. Hammond & Jared Diamond

Why are sustained energy budgets of humans and other vertebrates limited to not more than about seven times resting metabolic rate? The answer to this question has potential applications to growth rates, foraging ecology, biogeography, plant metabolism, burn patients and sports medicine.

The highest measured daily energy budgets maintained by any humans over long periods were achieved in the 1984 Tour de France race¹. For 22 days the world's best cyclists, in peak physical condition and fanatically motivated, pedalled up and down 34 mountains and over 3,826 km. Metabolic studies of four of those who completed the race showed that they expended energy at an average rate of 7,000 kcal per day. Nevertheless, throughout the day they gulped so many energy-rich drinks, and at night they bolted down so much pasta, bread and butter, and sweet cakes, that their body mass and fat content at the race's conclusion were virtually the same as at the start. That is, they fuelled their energy expenditure by concurrent energy intake, not by depleting body energy reserves.

To most of us 'couch potatoes', 7,000 kcal d⁻¹ sounds impressive. It is about 4.3 times the basal metabolic rate (BMR) of 1,640 kcal d⁻¹ expected for 70-kg men such as these athletes, BMR being the energy budget of someone fasting and lying motionless in bed all day and night. Actual energy budgets for sedentary men—such as typical laboratory scientists, bankers and *Nature* readers—are barely 70% higher than BMR, namely 2,800 kcal d⁻¹. Budgets are still only 3,800 kcal d⁻¹ even for miners, 4,100 kcal d⁻¹ for army recruits in training, and 5,000 kcal d⁻¹ for Robert Scott and his ill-fated companions man-hauling heavy sleds up glaciers in the Antarctic cold².

Yet, rather than being impressed by the Tour de France cyclists' high energy budgets, we have to wonder why they were not even higher. If any of those athletes had been able to digest 17,000 kcal d⁻¹ instead of 7,000, and to convert it to muscular work, he would have left his competitors behind in the dust. The fact that even such motivated fanatics did not achieve such high budgets strongly suggests that they were incapable of doing so. Are humans really restricted to time-averaged energy budgets below about five times BMR? If so, why?

Energy budgets of animals

Sustained metabolic rate (SusMR) may be defined as the time-averaged energy budget that an animal or human maintains over times sufficiently long that body mass remains constant because time-averaged energy intake equals time-averaged energy expenditure. Much higher metabolic rates can of course be achieved over short times (for example, the time required for a 200-m dash), but they are fuelled by consumption of body energy reserves and not sustained by concurrent energy intake.

Table 1 lists measurements of daily energy budgets (SusMR) in 50 active wild vertebrate species. To normalize for species differences in body size and also in basal energy expenditure, SusMR is divided by resting metabolic rate (RMR) to obtain the ratio SusMR/RMR, termed sustained metabolic scope (SusMS). (RMR would equal BMR if the animal were fasting and not growing or producing new biomass, but these conditions cannot automatically be assumed to apply to a wild-caught animal.) Measured metabolic scopes of active wild animals fall in the range 1.3 to 7.0. Of the 50 species studied, 11 achieved scopes modestly exceeding that of Tour de France cyclists,

but no scope measured to date exceeds 7.0. Yet some of these scope measurements apply to animals that one might expect to have been maximally motivated, just as are Tour de France cyclists: for example, birds foraging from dawn to dusk to gather food for their nestlings, or small mammals nursing pups or active at low environmental temperatures. Thus, Table 1 suggests that energy budgets of wild animals may be constrained by a ceiling similar to that limiting human athletes³. What factors could impose that ceiling?

Factors that might impose metabolic ceilings

If one accepts that sustained metabolic ceilings exist, four hypotheses about the responsible proximate factors suggest themselves.

First, ceilings might be imposed by food availability and not by any property of animals' bodies. This hypothesis is difficult to exclude in studies of wild animals, whose available food supply is generally unmeasurable: perhaps their budgets would have been higher if more food had been available. However, this hypothesis cannot apply to Tour de France cyclists, who certainly have access to unlimited food.

If the limit resides in properties of animals' bodies rather than in the food supply, one can then distinguish two further hypotheses: so-called peripheral limitations associated with energy-consuming machinery, and so-called central limitations associated with energy-supplying machinery. As for the former (our second hypothesis), energy budgets might be limited during physical exercise by work properties of skeletal muscle, during lactation by the milk-producing capacity of mammary glands or by availability of limiting nutrients, and during cold exposure by the capacities of brown fat or muscle shivering for heat production. Perhaps the gut and other central energy-supplying organs could supply calories faster, but those peripheral organs cannot (supposedly) convert the calories to work and heat faster. If so, one might expect in a given animal species different values of the metabolic ceiling for different modes of energy expenditure, depending on the particular properties of skeletal muscle or mammary glands or brown fat in that species.

The third hypothesis conversely attributes metabolic ceilings to bottlenecks in shared central machinery for acquiring, processing and distributing calories to any and all energy-consuming organs⁴. For example, the bottleneck could reside in the gut's capacity to digest and absorb food, or in the liver's capacity to process the absorbed foodstuffs, or in the lungs' capacity to take up O₂ or exhale CO₂, or in the heart's capacity to pump around the O₂ or nutrients, or in the kidneys' capacity to excrete the resulting wastes. Perhaps muscles could work faster and mammary glands could secrete milk faster, if only the rest of the body could supply nutrients and remove wastes faster. If such central limits applied, a given animal species might exhibit the same value of metabolic ceiling for different modes of energy expenditure.

Finally, animals might have evolved such that capacities of several of these potential limiting factors are matched to each other, a possibility termed symmorphosis⁵. For instance, the maximal rates

Table 1 The highest values of sustained metabolic scope (SusMS) reported for 50 vertebrate species

Species	Condition	Body mass (g)	RMR (kJ g ⁻¹ d ⁻¹)	SusMR (kJ g ⁻¹ d ⁻¹)	SusMS	Reference
MAMMALS						
Rodents						
<i>Acomys cahirinus</i>	Lactating (3 pups)	40	0.49	3.25	6.7	30
<i>Acomys russatus</i>		51	0.37	1.07	2.7	2
<i>Ammospermophilus leucurus</i>		96	0.49	0.99–1.43	1.7–2.6	2
<i>Apodemus flavicollis</i>	Cold (–10°C)	28.4	1.13	4.16	3.7	31
<i>Microtus agrestis</i>	Cold (5°C)	23.7	2.34	6.68	2.9	32
<i>Mus musculus</i>	Lactating (14 pups)	38.2	1.16	7.57	6.5	K.A.H. (unpubl.)
<i>Peromyscus leucopus</i>		21	0.76	3.28–4.28	3.9–5.0	2
<i>Peromyscus maniculatus</i>	Cold field measurements	20.6	0.42	2.72	6.4	33
<i>Sekeetamys calurus</i>		41	0.37	1.07	2.9	2
<i>Spermophilus saturatus</i>	Lactating (5 pups in the field)	232	0.43	1.59	3.7	34
<i>Thomomys bottae</i>		143	0.40	1.28–1.37	2.8–3.0	2
Marsupials						
<i>Antechinus stuartii</i>		37	0.46	2.72	4.6	2
<i>Gymnobelideus leadbeateri</i>		166	0.29	1.87–1.98	5.6–5.8	2
<i>Macropus eugenii</i>		4,800	0.14	0.27	1.9	2
<i>Petaurus breviceps</i>		128	0.32	1.37–1.71	4.0–4.4	2
<i>Phascolarctos cinereus</i>		4,770	0.20	0.26	2.3–3.0	2
<i>Pseudocheirus peregrinus</i>		860	0.16	0.78	4.7	2
<i>Setonix brachyurus</i>		2,510	0.15	0.32–0.44	1.8–1.9	2
<i>Sminthopsis crassicaudata</i>		14	0.64	4.06	6.9	2
Other mammals						
<i>Alouatta palliata</i>		4,670	0.21	0.35–0.89	1.5–2.0	2
<i>Bradypus variegatus</i>		3,790	0.09	0.13–0.19	1.4–2.0	2
<i>Echinops telfairi</i>	Lactating (1 pup)	262.6	0.11	0.23	2.2	35
<i>Homo sapiens</i>	Tour de France cyclist	68,500	0.10	0.42–0.56	42–5.6	2
BIRDS						
<i>Aethia pusilla</i>	Rearing chicks	83	1.38	4.32	3.1	36
<i>Alectoris chukar</i>		475	0.36	0.65–1.16	1.6–2.4	2
<i>Callipepla gambelii</i>		126	0.52	0.63	1.3	2
<i>Delichon urbica</i>		21	1.48	3.90–4.21	2.6–2.8	2
<i>Diomedea exulans</i>		8,130	0.22	0.36–0.54	1.6–2.0	2
<i>Eudypetes chrysolophus</i>		3,870	0.19	1.12	5.8	2
<i>Falco tinnunculus</i>		108	0.66	1.56	2.9	2
<i>Hirundo tahitica</i>	Rearing chicks	14	1.09	4.25	3.9	37
<i>Macronectes giganteus</i>		4,780	0.24	1.16–1.32	4.3–4.4	2
<i>Merops viridis</i>	Rearing chicks	34	0.76	2.55	3.4	2
<i>Oceanites oceanicus</i>	Rearing chicks	42.2	0.87	3.72	4.3	38
<i>Oenanthe oenanthe</i>	Rearing chicks	24	1.54	3.67	2.4	10
<i>Pelecanoides urinatrix</i>	Rearing chicks	132	0.96	4.28	4.5	10
<i>Pelecanoides georgicus</i>	Rearing chicks	119	1.02	3.89	3.9	10
<i>Puffinus pacificus</i>		338	0.38	1.60	4.3	2
<i>Pygoscelis adeliae</i>		3,970	0.27	1.03	3.8	2
<i>Pygoscelis papua</i>		6,290	0.26	0.59–0.64	2.3–2.5	2
<i>Rissa tridactyla</i>		386	0.81	2.27	2.8	2
<i>Sterna fuscata</i>		148	0.47	1.30	3.0	2
<i>Sula bassanus</i>	Rearing chicks	3,210	0.23	1.52	6.6	39
<i>Sula capensis</i>	Rearing chicks	2,575	0.32	1.31	4.7	40
<i>Sula sula</i>	Breeding and free-ranging	1,070	0.78	1.21	1.6	41
REPTILES						
<i>Amblyrhynchus cristatus</i>		965–2,250	0.04–0.08	0.06–0.13	1.7	2
<i>Cnemidophorus tigris</i>		16	0.13	0.21	1.6	2
<i>Cnemidophorus hyperythrus</i>		3.9–4.4	0.17–0.21	0.22–0.37	1.3–1.8	2
<i>Dipsosaurus dorsalis</i>		3.5–57	0.13–1.15	0.21–2.34	1.0–2.6	2
<i>Sceloporus virgatus</i>		5.5–76	0.04–0.08	0.13–0.21	2.1–4.6	2

SusMS is calculated as the ratio SusMR/RMR. Most values of SusMR and RMR were measured in the same individual animals, except for some of the values from ref. 2. The condition of the animals is listed if it is known.

Table 2 Sustainable metabolic scopes (SusMS), measured in the same species under several different types of energy demand

Taxa	Condition	Body mass (g)	RMR ($\text{kJ g}^{-1} \text{d}^{-1}$)	SusMR ($\text{kJ g}^{-1} \text{d}^{-1}$)	SusMS	Reference
MAMMALS						
Rodents						
<i>Acomys cahirinus</i>	Non-lactating, non-cold	42	0.52	1.37	2.5	2
	Cold (3°C)	37.8	0.49	1.845	3.8	30
	Lactating (3 pups)	40	0.49	3.25	6.7	30
	Cold (-10°C)	28.4	1.13	4.16	3.7	30
<i>Peromyscus maniculatus</i>	Lactating (3–5 pups)	24.7	0.88	4.05	4.6	42
	Cold (-10°C)	21.4	0.90	5.45	6.1	43
<i>Mus musculus</i>	Non-pregnant, non-lactating	30.8	0.79	1.81	2.3	8
	Strenuous exercise (running 6 h d^{-1})	25.2	0.81	2.94	3.6	I. Choshniak <i>et al.</i> (unpublished results)
	Cold (5°C)	30.7	0.90	3.59	4.0	K.A.H. (unpublished results)
	Cold (-10°C)	26.9	1.06	5.09	4.8	7
	Lactating (5 pups)	35.9	0.94	5.77	6.1	K.A.H. (unpublished results)
	Lactating (14 pups)	38.2	1.16	7.57	6.5	K.A.H. (unpublished results)
<i>Humans</i>	Lactating (14 pups) in cold (5°C)	39.3	1.27	8.33	6.6	K.A.H. (unpublished results)
Nepali women Nov–Dec	Light physical activity	46,670	0.11	0.18	1.6	44
Nepali women July–Sept	Pregnancy and heavy physical activity	52,000	0.10	0.19	1.9	44
	Lactation and heavy physical activity	47,900	0.11	0.20	1.9	44
	Very heavy physical activity	46,670	0.11	0.22	2.0	44
BIRDS						
<i>Oceanites oceanicus</i>	Incubation and brooding of eggs	42.2	0.87	2.81	3.2	38
	Rearing chick and at sea	42.2	0.87	3.72	4.3	38

at which the gut could absorb calories, the lungs could take up the O_2 necessary to burn them aerobically, and the muscles could convert the calories to work might have evolved to be about the same.

Experimental tests of ceilings in animals

A key step in confirming the suggested reality of metabolic ceilings and distinguishing between hypotheses about their origin has been laboratory studies in which animals, in the presence of excess quantities of available food, have been pushed to maximal SusMRs through different modes of energy expenditure. These studies serve three purposes: they exclude food availability as a potential limitation; they test whether rates of energy expenditure in fact increase up to a ceiling or continue to increase with increasing demands; and they test whether, within a given species, that ceiling value (if it exists) differs among different modes of energy expenditure, as predicted by the peripheral limitation hypothesis but not by the central limitation hypothesis. In effect, these experiments seek to study animals under conditions of maximal motivation, similar to those prevailing for Tour de France cyclists.

These studies do reveal maximal energy budgets with increasing 'motivation'. For instance, laboratory mice held at lowered ambient temperature counter their increased heat loss with increased heat production by non-shivering thermogenesis, yet maintain body mass by correspondingly increasing their food intake to fuel that increased metabolic rate^{6,7}. Their food intake increases with decreasing ambient temperature down to -15°C , below which they lose body mass and die in the presence of excess food. Thus, even given the maximal possible motivation of saving their lives, they do not

increase food intake and heat production further. Similarly, mice motivated by motherly love increase their food intake 4.9-fold to produce milk for a normal litter of 8–10 pups⁸. They increase food intake and milk output slightly more as motivation is experimentally increased by cross-fostering additional pups from another litter of the same age, or by experimentally requiring the mother to nurse her pups beyond the usual age of pup weaning until pups weigh double their usual mass at weaning^{8,9}. But food intake with either of these two experimentally increased appeals to motherly love still reaches a ceiling, in the presence of excess food, of 5.5 times non-lactating food intake. Finally, mice required to run on treadmills increase their food intake and run for up to 6 h d^{-1} , after which they become exhausted and run no further despite the availability of excess food (I. Choshniak *et al.*, unpublished observations).

Table 2 depicts the results of such studies for five vertebrate species in which ceilings were sought experimentally by more than one mode of energy expenditure. These results yield the following conclusions.

(1) Ceiling values of the ratio SusMS fall in the range of 4.3–6.7 (Table 2), no higher than values measured in the field for some wild animals (Table 1). Hence some of the wild animals studied in Table 1 were also probably at or near their metabolic ceilings.

(2) In each animal species, the ceiling value differs for different modes of energy expenditure. For instance, in laboratory mice peak SusMS is 3.6 for physical exercise, 4.8 for heat production at low ambient temperature, and 6.5 for lactation. This agrees with the peripheral limitation hypothesis and disagrees with the central limitation hypothesis, for all those modes whose scope value is surpassed during some other mode in the same species (such as

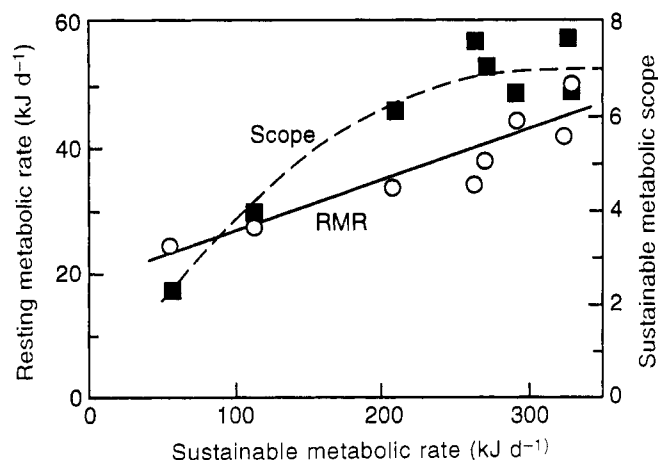


Figure 1 Measured or estimated values of RMR, and calculated values of SusMS (metabolic scope), in female mice, plotted as a function of SusMR. SusMR was manipulated experimentally by applying cold stress, lactation with various litter sizes from 5 to 14 pups, or cold stress plus lactation. The left-most point is for virgin mice at 23°C; the right-most point, for mice nursing 14 pups at 5°C. Note that both RMR and SusMS increase with SusMR. The straight line through RMR values, and the curve fitted by eye through SusMS values, serve merely to summarize trends and have no theoretical significance. Data are from ref. 8 and our unpublished measurements.

exercise and heat production in laboratory mice). However, the central limitation hypothesis is not thereby excluded for the mode yielding the highest scope value in a given species (such as lactation in laboratory mice).

(3) Different species excel for different modes of energy expenditure. For instance, the highest values of SusMS are achieved during lactation by laboratory mice, at low ambient temperatures for a wild mouse population (*Peromyscus maniculatus*), and during physical exercise for humans. These differences are obviously related to each species' lifestyle: the wild mouse populations studied live without clothing in cold environments, and traditional human populations were physically vigorous but did wear clothing in cold environments and gave birth to a litter of only one, whereas laboratory mice have been selected for the ability to rear large litters.

RMR increases at high SusMR

A clue as to why SusMS reaches a ceiling emerged from a study applying a combination of two energy stresses (cold and lactation) to laboratory mice⁹. Food intake and SusMS reached an upper limit during maximal tolerated cold stress, and reached a higher limit during maximal lactation. But when cold stress and lactation were applied simultaneously, food intake increased even further (by 16%) above the lactational ceiling, by an absolute increment roughly equal to that measured for cold exposure of non-lactating mice. Evidently, even though food intake did not increase any further under the stress of lactation alone, lactating mice were still capable of processing more food energy and expending it for a different purpose. Nevertheless, SusMS (= SusMR/RMR) was no higher in cold lactating mice than in lactating mice at room temperature, because RMR as well as SusMR proved to be higher in cold lactating mice.

Figure 1 illustrates that RMR and SusMS of mice increase with increases of SusMR elicited experimentally by any means. As SusMR increases sixfold, RMR increases twofold, and SusMS reaches values near seven. These results suggest that SusMR and RMR are obligatorily somehow coupled to each other, in such a way that values of their ratio are highest at high SusMR (Fig. 1).

One test of this hypothesis has involved analysing SusMR and RMR values for a wide variety of mammal and bird species under

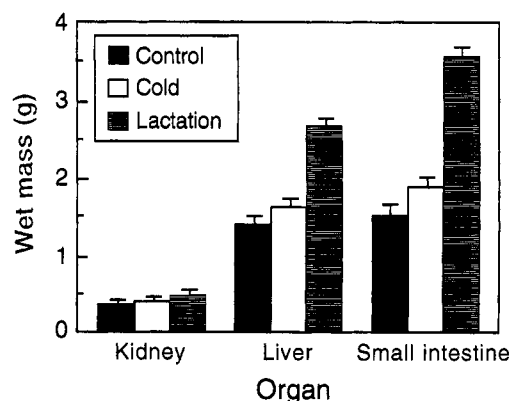


Figure 2 Changes in wet mass of kidney, liver and small intestine in mice with three increasing levels of energy demand (from left to right: control or virgin females, acclimation to 5°C, and lactation with a litter of 8 pups). Within each organ, changes of organ mass with energy demand are statistically significant at $P < 0.05$ by ANOVA. Data are from ref. 8.

a wide variety of conditions (see for example, refs 10–12). Perhaps not surprisingly, those analyses have yielded at best equivocal evidence for a coupling between SusMR and RMR, because of likely species differences in the form of the relation and because of the modest range of mass-specific SusMR and RMR values observed for different mammal and bird species of a given body size. However, the hypothesis is supported by the much greater contrast between endothermic and ectothermic vertebrate species: endotherms (mammals and birds) have both SusMR values and RMR values an order of magnitude higher than those of ectotherms (reptiles, amphibia and fish) of the same body size (see for example, refs 13–15).

Mechanism of increased RMR at high SusMR

Although Fig. 1 suggests a formal explanation for an upper asymptote of SusMS, it leaves unanswered the question of the mechanism underlying the observed increase in RMR with increasing SusMR. We suggest that that mechanism depends in large part on mass increases in energy-supplying organs.

As is well known and obvious, the masses of peripheral energy-consuming organs increased selectively and reversibly during conditions of high sustained function. For instance, specific skeletal muscles can be observed to grow in athletes during training; mammary glands grow during lactation; and both organs atrophy again upon resumption of couch-potato or non-lactating status. But there is also, taking place invisibly inside an animal's body during any such condition of high SusMR, a less well-known reversible hypertrophy of central energy-supplying organs. In laboratory mice, the small intestine, kidneys, liver and heart increase in mass during lactation, cold exposure, or both stresses applied simultaneously^{6–9,16} (Fig. 2). These same organs, plus the lungs, stomach and pancreas, increase in mass in snakes and frogs with greatly elevated metabolic rates during digestion of large meals^{17,18}. The reason why these organs must grow under conditions of high food intake and energy budgets is that their masses under resting conditions provide only modest functional reserve capacities that would be unequal to the demands upon them at high SusMR if their masses did not increase^{8,9,16}.

Analogous to these reversible intraspecific mass increases in

energy-supplying organs at high SusMR are the permanent inter-specific mass differences observed in the same organs when one compares species with chronically high or low energy budgets. The largest contrast is between endotherms and ectotherms: associated with their 15–50-fold higher SusMRs, endotherms have intestines, livers, kidneys and hearts several times larger than those of ectotherms of the same body size^{13,15,19}. Temperate-zone birds and mammals tend to have larger hearts, kidneys, livers, intestines, stomachs and pancreases than tropical birds and mammals, which tend to operate at lower SusMRs because of less need for heat production²⁰. Among different species of temperate-zone birds, those with higher SusMRs and RMRs tend to have larger hearts and kidneys¹⁰. If RMR differences may be taken as an indication of corresponding SusMR differences, it is also relevant to note that, among different strains of laboratory mice, those with higher RMRs tend to have larger hearts, kidneys, small intestines and livers²¹.

These central energy-supplying organs, although their masses contribute only a modest fraction of an animal's total body mass, contribute a large fraction of an animal's RMR because of their very high mass-specific metabolic rates. For instance, in laboratory mice the kidneys have a mass-specific metabolic rate 51% higher than that of any other tissue, and the heart, small intestine and liver are not far behind²². Even in a resting animal or human, the kidney has a very high metabolic rate because it is constantly filtering urine and then actively reabsorbing most of the filtrate; the small intestine has a high metabolic rate because its mass turns over every 2–3 d; and the liver is constantly processing nutrients. Although these three organs contribute, respectively, only 0.5, 5 and 7% of a mouse's body mass, in aggregate they account for 24% of RMR²². As another example, the liver and gut account for up to 70% of the metabolic heat production above maintenance in ruminants experiencing high workloads²³. This, too, is a large fraction of the metabolic heat production from a relatively small fraction of the total body mass.

Thus, our suggested mechanistic explanation of Fig. 1, and of the observed ceiling on SusMS, is that high SusMR requires increased mass of energy-supplying organs, but that the high maintenance and operating costs of these organs cause them to contribute disproportionately to RMR, which thus tends to increase with SusMR. That is, high energy budgets depend on energetically expensive metabolic machinery.

Future directions

We conclude by pointing out some questions ripe for study.

Relationships between potential limiting factors. Depending on the circumstances and on the mode of energy expenditure, SusMR might be limited by the energy-consuming organ involved or by one or more energy-supplying organs. But it might alternatively be limited by several such organs in concert. That is, if natural selection has acted to minimize unutilized reserve capacities, then one might observe a rough match between the intestine's capacity to digest and absorb nutrients, the liver's capacity to process absorbed nutrients, the lungs' capacity to take up the O₂ necessary for their metabolism, the heart's capacity to pump the O₂ and nutrients, the capacity of the highest-capacity energy-consuming organ to utilize the nutrients, and the kidneys' capacity to excrete the resulting wastes. Taylor and Weibel³ refer to this possibility as *symmorphosis*. On the other hand, there are also plausible reasons why natural selection might have led to a different outcome. This central question of evolutionary design is unresolved and in need of study²⁴.

Ecological, biogeographic and behavioural significance of ceilings on SusMS. Physiological ceilings on SusMS imply ceilings on heat production, which might account for observed distributional limits of some species in cold climates (poleward latitudinal boundaries and upper altitudinal boundaries)²⁵. Physiological ceilings may also constrain daily foraging efforts, lactational performance and other life-history attributes.

Interspecific variation in ceilings. One might guess that peak values of SusMS differ among wild animal species: for example, that values may be higher for wolves than for sloths. One could also speculate whether species of relatively constant mesic environments should tend to have higher or lower ceilings than species of unpredictable environments. The data of Table 1 hint at such differences but are equivocal, because many SusMS values in Table 1 may not represent peak values. Comparative studies are badly needed that push numerous species to peaks of SusMS.

Applications to growth. Biosynthesis for growth represents a form of sustained energy expenditure and should therefore be subject to a ceiling that could be compared with ceilings on other forms of energy expenditure. The main practical problem in doing so is to know how to divide the measured energy budget of growing animals into the cost of growth and the cost of maintenance activities. Several experimental solutions to this problem have been proposed, such as restricting food to halt growth, estimating costs of protein synthesis, and administering inhibitors of protein synthesis (see for example ref. 26).

Applications to other forms of biomass production. During physical exercise or heat production, dietary nutrients are ingested, absorbed, and then combusted and converted to work or heat, and the combustion is reflected in O₂ consumption and CO₂ production. However, during biomass production, such as lactation, egg production, fat storage, or somatic growth, some of the ingested and absorbed nutrients are not combusted but are transformed into the nutrients of the produced biomass. Hence daily energy budgets calculated from food assimilation should equal budgets calculated from O₂ consumption or CO₂ production during exercise or heat production but might exceed the latter budgets during biomass production, to a degree depending on costs of production and the metabolic transformations involved. Is such a discrepancy actually observed? If intestinal digestion and absorption were limiting, food intake might be no higher during peak biomass production than during peak exercise or heat production; is that true?

Applications to plants. Can one measure, in photosynthetically active plants, quantities that reflect maintenance costs and total daily energy budget and that are, respectively, analogous to RMR and SusMR of animals? For instance, is dark respiration rate or maintenance respiration of plants analogous to animals' RMR, and growth respiration or total respiration of plants analogous to SusMR? Are these two quantities loosely coupled in plants as they are in animals? Is that coupling reflected in the observed correlation between plant growth rates and respiration rates or calorimetrically measured metabolic heat rates^{27,28}?

Clinical applications. People who have suffered burns over more than 40% of their body surface exhibit large increases in resting metabolic rate (termed 'hypermetabolism') because of two ongoing costly processes: tissue repair and heat production to balance increased heat loss²⁹. Burn patients often receive nutrients intravenously by parenteral nutrition. To speed their healing, their metabolic rate should be as high as possible. What limits their metabolic rate and their capacity to process parenterally administered nutrients?

Sports medicine. The Tour de France is unusual among sports events in that it lasts so long that energy expenditure over the course of the event must be balanced in large part by concurrent energy intake. For any event lasting less than 1 d, energy expenditure instead comes largely from depletion of energy reserves not replenished until after the event. Hence SusMS ceilings are not expected to limit performance in a short event but should certainly limit athletes' training regimes leading up to the event. Are human training regimes really limited by properties of skeletal muscle, as usually assumed, or instead by energy-supplying organs such as the intestine and kidneys? Could knowledge of such limitations help in improving training regimes? Heart enlargement in athletes is well known, but one might also expect enlargement of

athletes' intestines, kidneys and livers, and consequent increase in BMR, by analogy with lactating and cold-exposed animals; is that true? □

Kimberly A. Hammond is in the Biology Department, University of California, Riverside, California 92521, and Jared Diamond is in the Physiology Department, University of California Medical School, Los Angeles, California 90095-1751, USA.

1. Westerterp, K. R. *et al.* Use of the doubly labeled water technique in humans during heavy sustained exercise. *J. Appl. Physiol.* **61**, 2162–2167 (1986).
2. Petersen, C. C. *et al.* Sustained metabolic scope. *Proc. Natl Acad. Sci. USA* **87**, 2324–2328 (1990).
3. Drent, R. H. & Daan, S. The prudent parent: energetic adjustments in avian breeding. *Ardea* **68**, 225–252 (1980).
4. Weiner, J. Metabolic constraints to mammalian energy budgets. *Acta Theriol.* **34**, 3–35 (1989).
5. Taylor, C. R. & Weibel, E. R. Design of the mammalian respiratory system. I. Problems and strategy. *Resp. Physiol.* **44**, 1–10 (1981).
6. Toloza, E. *et al.* Nutrient extraction by cold-exposed mice: a test of digestive safety margin. *Am. J. Physiol.* **261**, G608–G620 (1991).
7. Konarzewski, M. & Diamond, J. M. Peak sustained metabolic rate and its individual variation in cold-stressed mice. *Physiol. Zool.* **67**, 1186–1212 (1994).
8. Hammond, K. A. *et al.* Metabolic ceilings under a combination of peak energy demands. *Physiol. Zool.* **68**, 1479–1506 (1994).
9. Hammond, K. A. & Diamond, J. M. Limits to dietary nutrient intakes and intestinal nutrient uptakes in lactating mice. *Physiol. Zool.* **67**, 282–303 (1994).
10. Daan, S. *et al.* Avian basal metabolic rates: their association with body composition and energy expenditure in nature. *Am. J. Physiol.* **259**, R333–R340 (1990).
11. Koteja, P. On the relation between basal and field metabolic rates in birds and mammals. *Funct. Ecol.* **5**, 56–64 (1991).
12. Ricklefs, R. E. *et al.* The relationship between basal metabolic rate and daily energy expenditure in birds and mammals. *Am. Nat.* **147**, 1047–1071 (1996).
13. Else, P. L. & Hulbert, A. J. Comparison of the 'mammal machine' and the 'reptile machine': energy production. *Am. J. Physiol.* **240**, R3–9 (1981).
14. Peters, R. H. *The Ecological Implications of Body Size* (Cambridge Univ. Press, Cambridge, 1983).
15. Nagy, K. A. Field metabolic rate and food requirement scaling in mammals and birds. *Ecol. Monogr.* **57**, 111–128 (1987).
16. Hammond, K. A. & Diamond, J. M. An experimental test for a ceiling on sustained metabolic rate in lactating mice. *Physiol. Zool.* **65**, 952–977 (1992).
17. Secor, S. M. *et al.* Rapid upregulation of snake intestine in response to feeding—a new model of intestinal adaptation. *Am. J. Physiol.* **266**, G695–G705 (1994).
18. Secor, S. M. & Diamond, J. M. Adaptive responses to feeding in Burmese pythons—pay before pumping. *J. Exp. Biol.* **198**, 1313–1325 (1995).
19. Karasov, W. H. *et al.* What transport adaptations enable mammals to absorb sugars and amino acids faster than reptiles? *Am. J. Physiol.* **249**, G271–G283 (1985).
20. Rensch, I. & Rensch, B. Relative Organmasse bei tropischen Warmblütern. *Zool. Anzeiger* **156**, 106–124 (1956).
21. Konarzewski, M. & Diamond, J. M. Evolution of basal metabolic rate and organ masses in laboratory mice. *Evolution* **49**, 1239–1248 (1995).
22. Martin, A. W. & Fuhrman, F. A. The relationship between summated tissue respiration and metabolic rate in the mouse and dog. *Physiol. Zool.* **28**, 18–34 (1955).
23. Johnson, D. E. *et al.* Changes in liver and gastrointestinal tract energy demands in response to physiological workload in ruminants. *J. Nutr.* **120**, 649–655 (1990).
24. Weibel, E. R. *et al.* *Diversity in Biological Design* (Cambridge Univ. Press, Cambridge, 1997).
25. Root, T. Energy constraints on avian distribution and abundances. *Ecology* **69**, 330–339 (1988).
26. Ricklefs, R. E. Adaptation, constraint, and compromise on avian post-natal development. *Biol. Rev.* **54**, 269–290 (1979).
27. Amthor, J. S. *Respiration and Crop Productivity* (Springer, New York, 1989).
28. Hansen, L. D. *et al.* Prediction of long-term growth rates of larch clones by calorimetric measurement of metabolic heat rates. *Can. J. For. Res.* **19**, 606–611 (1989).
29. Royall, D. *et al.* Continuous measurements of energy expenditure in ventilated burn patients: an analysis. *Critical Care Med.* **22**, 399–406 (1994).
30. Koteja, P. *et al.* Maximum cold and lactation induced rate of energy assimilation in *Acomys cahirinus*. *Polish Ecol. Studies* **20**, 369–374 (1994).
31. Koteja, P. Maximum cold-induced energy assimilation in a rodent, *Apodemus flavicollis*. *Comp. Biochem. Physiol.* **112**, 479–485 (1995).
32. McDevitt, R. M. & Speakman, J. R. Central limits to sustainable metabolic rate have no role in cold acclimation of the short-tailed field vole (*Microtus agrestis*). *Physiol. Zool.* **67**, 1117–1139 (1994).
33. Hayes, J. P. Altitudinal and seasonal effects on aerobic metabolism of deer mice. *J. Comp. Physiol.* **159**, 453–459 (1989).
34. Kenagy, G. J. *et al.* Energy expenditure during lactation in relation to litter size in free-living golden-mantled ground squirrels. *J. Anim. Ecol.* **59**, 73–88 (1990).
35. Poppitt, S. D. *et al.* Energetics of reproduction in the lesser hedgehog tenrec *Echinops telfairi* (Martin). *Physiol. Zool.* **67**, 976–994 (1994).
36. Roby, D. D. & Ricklefs, R. E. Energy expenditure in adult least auklets and diving petrels during the chick-rearing period. *Physiol. Zool.* **59**, 661–678 (1986).
37. Bryant, D. C. *et al.* Reproductive energetics of two tropical bird species. *Auk* **101**, 25–37 (1984).
38. Obst, B. S. *et al.* Energy utilization by Wilson's storm-petrels (*Oceanites oceanicus*). *Physiol. Zool.* **60**, 200–210 (1987).
39. Brit-Friesen, V. L. *et al.* Activity-specific metabolic rates of free-living northern gannets and other seabirds. *Ecology* **70**, 357–367 (1989).
40. Adams, N. J. *et al.* Energy expenditure and food consumption by breeding Cape gannets *Morus capensis*. *Mar. Ecol. Progr. Ser.* **70**, 1–9 (1991).
41. Ballance, L. T. Flight energetics of free-ranging red-footed boobies (*Sula sula*). *Physiol. Zool.* **68**, 887–914 (1995).
42. Koteja, P. Limits to the energy budget in a rodent, *Peromyscus maniculatus*: the central limitation hypothesis. *Physiol. Zool.* **69**, 981–993 (1996).
43. Koteja, P. Limits to the energy budget in a rodent, *Peromyscus maniculatus*: does gut capacity set the limit? *Physiol. Zool.* **69**, 994–1020 (1996).
44. Panter-Brick, C. Seasonality of energy expenditure during pregnancy and lactation for rural Nepali women. *Am. J. Clin. Nutr.* **57**, 620–628 (1993).

Acknowledgements. We thank M. Chappell, K. Nagy and T. O'Connor for comments on the manuscript, L. Hansen, R. Pearcy and P. Nobel for discussion of applications to plants, and the National Institutes of Health for support.

KNOW YOUR COPYRIGHT RESPECT OURS

The publication you are reading is protected by copyright law.
Photocopying copyright material without permission is no different from stealing a magazine from a newsagent, only it doesn't seem like theft.

If you take photocopies from books, magazines and periodicals at work your employer should be licensed with CLA.
Make sure you are protected by a photocopying licence.



The Copyright Licensing Agency Limited
90 Tottenham Court Road, London W1P 0LP
Telephone: 0171 436 5931 Fax: 0171 436 3986

Structure of 20S proteasome from yeast at 2.4 Å resolution

Michael Groll, Lars Ditzel, Jan Löwe*, Daniela Stock*, Matthias Bochtler, Hans D. Bartunik† & Robert Huber§

Max-Planck-Institut für Biochemie, D-82152 Martinsried, Germany

† Arbeitsgruppen für strukturelle Molekularbiologie der Max-Planck-Gesellschaft, Notkestr. 85, D-22603 Hamburg, Germany

The crystal structure of the 20S proteasome from the yeast *Saccharomyces cerevisiae* shows that its 28 protein subunits are arranged as an $(\alpha 1 \dots \alpha 7, \beta 1 \dots \beta 7)_2$ complex in four stacked rings and occupy unique locations. The interior of the particle, which harbours the active sites, is only accessible by some very narrow side entrances. The β -type subunits are synthesized as propeptides before being proteolytically processed for assembly into the particle. The proforms of three of the seven different β -type subunits, $\beta 1$ /PRE3, $\beta 2$ /PUP1 and $\beta 5$ /PRE2, are cleaved between the threonine at position 1 and the last glycine of the pro-sequence, with release of the active-site residue Thr 1. These three β -type subunits have inhibitor-binding sites, indicating that PRE2 has a chymotrypsin-like and a trypsin-like activity and that PRE3 has peptidylglutamyl peptide hydrolytic specificity. Other β -type subunits are processed to an intermediate form, indicating that an additional nonspecific endopeptidase activity may exist which is important for peptide hydrolysis and for the generation of ligands for class I molecules of the major histocompatibility complex.

The proteasome or multicatalytic protease complex is the central enzyme of protein degradation in both the cytosol and the nucleus. It is involved in many biological processes, including the removal of abnormal, misfolded or improperly assembled proteins, stress response (by processing or degrading transcriptional regulators) cell-cycle control (by degradation of cyclins), cell differentiation and metabolic adaptation (by destruction of transcription factors or metabolic enzymes), and the cellular immune response (by generating antigenic peptides presented by major histocompatibility complex (MHC) class I molecules). These cellular functions are linked to a ubiquitin- and ATP-requiring protein degradation pathway involving the 26S proteasome, whose core and proteolytic chamber is formed by the 20S proteasome. Its essential nature is reflected in the lethality of chromosomal deletions of all but one ($\alpha 3$ /Y13; ref. 1) of its subunits in yeast^{2–13}. Proteasomes are ubiquitous in all three kingdoms of life. The 20S proteasome from the archaeobacterium *Thermoplasma acidophilum* has been analysed by X-ray crystallography at 3.4 Å resolution¹⁴. It has a cylindrical shape, with overall dimensions of 148 Å length and of 113 Å maximum and 75 Å minimum diameter, respectively. It is composed of 28 subunits arranged in a particle as four homoheptameric rings $\alpha_7\beta_7\beta_7\alpha_7$ with *D*₇ symmetry¹⁴. In the *T. acidophilum* proteasome, the amino-terminal threonine of the β -subunits is the binding site of inhibitory peptide aldehydes and is essential for its hydrolytic activity^{14,15}, which relates it to the Ntn ('N-terminal nucleophile') family of hydrolases¹⁶.

The yeast *Saccharomyces cerevisiae* 20S proteasome is similar in size and shape to the archaeobacterial proteasome but is much more complex, having seven different α -type and seven different β -type subunits, all of which have been cloned and sequenced and can be grouped by sequence homology¹⁷. As defined by the character of the P1 cleavage sites of chromogenic reporter groups, trypsin-like (basic), chymotrypsin-like (hydrophobic), and peptidylglutamyl-peptide hydrolytic (PGPH) (acidic) activities have been ascribed to the eukaryotic proteasome. These specificities, however, are not reflected in the cleavage-site pattern of its protein substrates, which are cleaved at almost every position^{18–21}.

Five of the seven β -type subunits are synthesized as propeptides, from which propeptides are cleaved during proteasome maturation^{11,15,22}. The role of these prosequences, apart from masking

the active sites before maturation, seems to differ between archaeobacteria and eukaryotes as the prosequences are not required in the former for proteasome assembly^{11,15,23–25}.

Eukaryotic 20S proteasomes, including those from yeast, are very closely related in their function, in the amino-acid sequences of their components and in their structure in the electronmicroscope. The α and β -type subunits of the mammalian 20S proteasome are arranged in an ordered and unique way²⁶. In mammalian cells, three additional non-essential subunits of the 20S proteasome, LMP2, LMP7 and MECL1, can replace constitutive components upon induction by the T-cell-derived antiviral cytokine interferon- γ . The LMP subunits are encoded by a region of the MHC gene cluster adjacent to the TAP ('transporters associated with antigen presentation') genes, which are involved in transport of antigenic peptides from the cytosol to the endoplasmic reticulum. Their expression or targeted deletion alters the proteasome's peptidase specificity and the amount of MHC class I cell-surface expression^{27–30}. The properties of peptide fragments generated by the proteasome correlate strongly with those of MHC class I ligands and processing by proteasomes is probably the principal pathway of MHC class I peptide generation¹².

Here we describe a crystallographic structural analysis of the yeast 20S proteasome. We have analysed crystals of proteasomes treated with a subunit-specific inhibitor (lactacystin), and with a non-specific tripeptide aldehyde inhibitor (calpain inhibitor I), respectively, at 2.4 Å resolution.

Crystallization and structure determination

An efficient preparation of the 20S proteasome from yeast was worked out, giving reproducible yields of ~50 mg crystallizable protein from 500 g yeast cells. The protein crystallizes from solutions containing Mg^{2+} and alcohol. The crystals are well ordered, but anisotropy leads to resolution of measurable reflections beyond 2.2 Å in the *b** direction and 2.8 Å perpendicular to *b**. A self-rotation function calculated at 5 Å resolution showed two crystal-symmetry-related peaks indicative of molecular local diad axes at $\Psi 86^\circ, \Phi 90^\circ$ and $\Psi 94^\circ, \Phi 90^\circ$. Their correlation values were one half of the value of the crystallographic diad, as predicted for almost ideal molecular 2-fold symmetry.

The *T. acidophilum* model was used for Patterson search calculations at 3.5 Å resolution. These showed a unique solution with a correlation value of 0.32 and an *R*-factor of 56% compared to the next

* Present address: MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK.

highest peak of 0.28 and 57%. The *T. acidophilum* model was truncated to polyaniline with only a few conserved residues retained in the α -type subunit. This model yielded an R -factor of 57.7% and was used to calculate a $2F_o - F_c$ map at 2.4 Å. The density was averaged using the 2-fold local axis of the model ($\Psi = 85.1$, $\Phi = 90.8$, $\kappa = 180.1$), backtransformed, and a new density was calculated with $2F_o - F_c$ coefficients. After 10 cycles of averaging, the quality of the map was good ($R_{\text{back}} = 27.3\%$) and showed little input model bias. The individual subunits were identified according to their characteristic amino-acid sequences and were built into the map. Anisotropy of diffraction was corrected in the last stages of refinement. The solvent-scattering contribution was included and taken into account during refinement and map calculation. In successive rounds of model building and refinement, the model was completed and included either lactacystin or the acetyl-Leu-Leu-norleucinal inhibitor molecules, 18 magnesium ions and 1,800 water molecules, in their respective complexes. Temperature factors were refined with restraints between bonded

atoms and between non-crystallographic-symmetry-related atoms. The R -values are satisfactory and the standard geometry of restrained bonds and angles and the local symmetry well maintained. The increase in R -value by 3% for data from 2.8 to 2.4 Å is a result of the anisotropic crystalline order causing a deterioration in data quality and of the very restricted inclusion of ordered solvent (Table 1).

Subunit structure

The fourteen genes cloned from yeast and shown to encode components of the 20S proteasome can be grouped into seven α -type and seven β -type subunits (see Methods and Table 2). The β -type subunits are synthesized as proproteins which are proteolytically processed to the mature forms found in the assembled particle. The proforms of $\beta 1/\text{PRE3}$, $\beta 2/\text{PUP1}$ and $\beta 5/\text{PRE2}$ are fully processed and cleaved between Gly -1 and Thr 1, with liberation of the active-site residue Thr 1. $\beta 7/\text{PRE4}$ is cleaved between Asn -9 and Thr -8 and $\beta 6/\text{C5}$ between His -10 and Gln -9 as stable

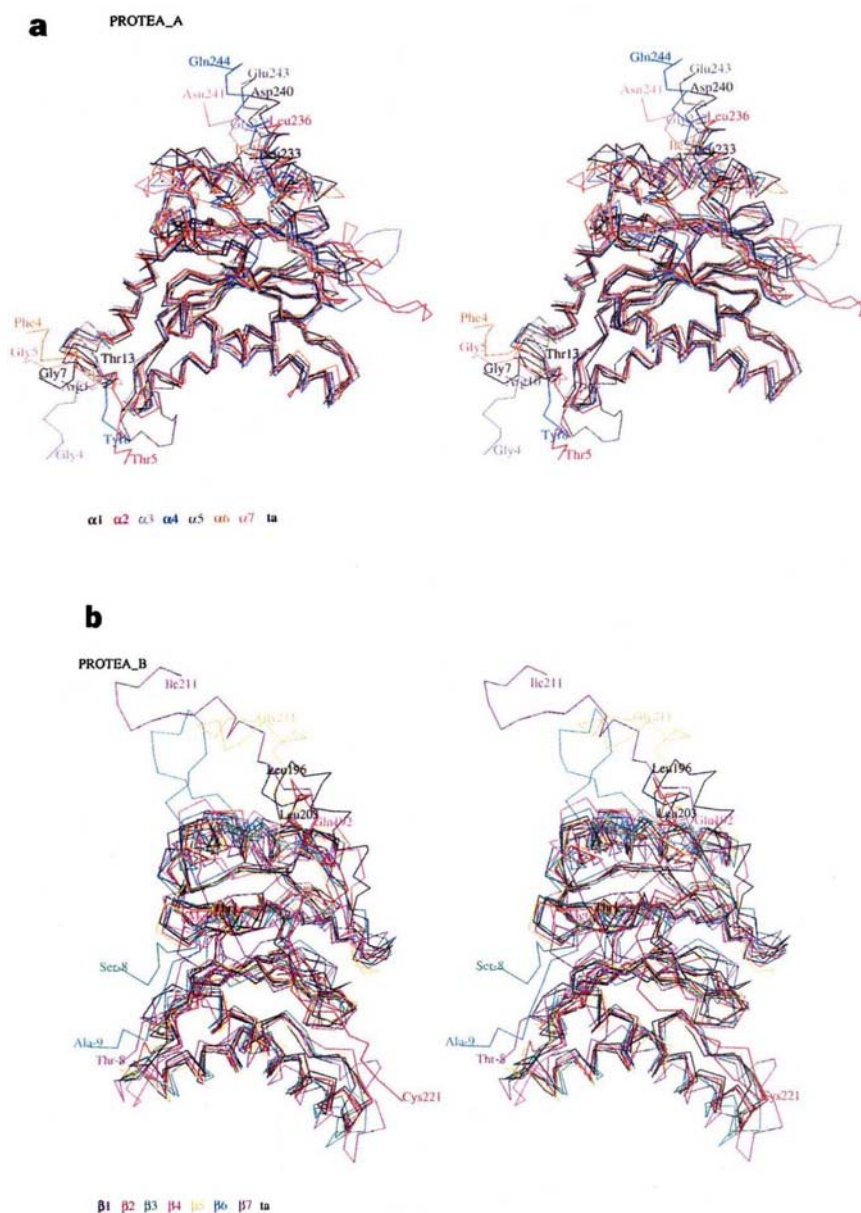


Figure 1 **a, b**, Overlay of the $C\alpha$ tracings of the seven α -type subunits of *S. cerevisiae* and of *T. acidophilum* (**a**) and of the β -type subunits (**b**). The orientations are similar to the setting in ref. 14.

processing intermediates. $\beta 4/C11$ and $\beta 3/PUP3$ are not processed from their primary translation products. All fourteen subunits are present in the molecular structure in unique locations. They are almost completely defined by electron density apart from some chain termini and one long insertion segment (Table 3). All seven α - and β -type polypeptides have the β -sandwich structure typical for Ntn-hydrolases. It is formed by two five-stranded antiparallel β -sheets, with helical layers on top made up of helices H3, H4, H5 and on the bottom by helices H1 and H2. Differences between the subunits appear in turns, which vary in length by one or two residues, in long insertions connecting secondary structural elements, and in the N-terminal and especially the C-terminal regions. The insertions extend existing secondary structures or form new ones which can be summarized as follows. The α -type subunit $\alpha 2/Y7$ has a long insertion loop between strands S9 and S10 composed of a short α -helix and two β -strands. $\alpha 4/PRE6$ and $\alpha 6/PRE5$ have longer helices H3 and H4. $\alpha 1/C7$ has the helix H3 extended by 2 turns. The subunits $\alpha 1/C7$, $\alpha 3/Y13$, $\alpha 4/PRE6$, $\alpha 5/PUP2$, $\alpha 7/C1$ have longer C-terminal helices H5 sticking out from the particle surface into solution. The highly charged, mostly acidic C-terminal segments are unstructured. The long insertions in the β -type subunits include an elaborate turn between helices H1 and H2 and an additional two-turn helix at residue 145 in $\beta 7/PRE4$ and a 17-residue insertion between H3 and H4, with a complex folding of strands and a short helix in $\beta 6/C5$. $\beta 2/PUP1$ has a very long C-terminal extension, which is disordered in its last 11 residues. Subunits $\beta 3/PUP3$ and $\beta 6/C5$ have short C termini, so helices H5 do not exist and the S10 strands are extended to enlarge the β -sheet. In $\beta 4/C11$, helix H5 exists but is shorter by two turns than in the

T. acidophilum model. The overlay of the alpha-carbon ($C\alpha$) tracings shows these insertions and also the considerable variability in the core structures, which is less in the α -type and more in the β -type subunits (Fig. 1a, b). Many of the subunit-specific turns, insertions, and N and C termini are involved in intersubunit contacts (see below).

The complex

The general architecture of the quarternary structure of the proteasome is the same in the *T. acidophilum* and yeast particles (Fig. 2a, b). The N-terminal loop segment, helix H0 (residues 20 to 30), loop L, the loop connecting H2 and S5, and strand S7 mediate α -*cis* interactions. The β -*cis* contacts, which seem to be less intimate, involve loop L, the N-end of helix H1, strand S7, and the turn connecting strand S8 and helix H3. These contacts are inherited

Table 1 Data collection and model refinement statistics

	Lactacystin	$P2_1$	Calpain inhibitor I
Space group			
Cell constants			
<i>a</i>	135.7 Å		135.7 Å
<i>b</i>	301.8 Å		302.3 Å
<i>c</i>	144.7 Å		150.0 Å
β	112.6°		111.5°
Limiting resolution	2.2 Å		2.3 Å
Measured reflections	1,191,158		885,201
Unique reflections	410,744		309,112
Completeness	96.5%		87.3%
R_{merge}^*	9.7%		14.2%
Reflections for refinement	402,338		299,618
Resolution range	50.0–2.4 Å		50.0–2.6 Å
Protein atoms		48,888	
Inhibitor atoms	30		162
Solvent molecules	1,800		742
$R_{\text{crist}}^\dagger$	23.0%	50–2.8 Å	26.0%
	26.0%	50–2.4 Å	27.7%
$R_{\text{free}}^\ddagger$	30.9%	50–2.4 Å	33.3%
$R_{\text{back}}^\ddagger$	13.0%	50–2.4 Å	19.4%
R.m.s. bond length §	0.010 Å		0.010 Å
R.m.s. bond angles §	1.8°		1.8°
R.m.s. NCS proteins §	0.13 Å		0.17 Å
R.m.s. NCS solvents §	0.17 Å		0.24 Å
Average <i>B</i>	31.1 Å ²		36.9 Å ²
R.m.s. bonded <i>B</i>	4.0 Å ²		3.3 Å ²
R.m.s. NCS <i>B</i> protein	2.0 Å ²		1.8 Å ²
R.m.s. NCS <i>B</i> solvent	3.3 Å ²		3.2 Å ²

* $R_{\text{merge}} = \sum |I - \langle I \rangle| / \sum I$, where *I* is the intensity value of an individual measurement and $\langle I \rangle$ is the corresponding mean value; the summation runs over all measurements.

$^\dagger R_{\text{crist}} = \sum ||F_{\text{obs}}| - |F_{\text{calc}}|| / \sum |F_{\text{obs}}|$, where $|F_{\text{obs}}|$ is the observed and $|F_{\text{calc}}|$ the calculated structure amplitude of a reflection. The summation is over all unique reflections. $R_{\text{back}} = \sum (||F_{\text{obs}}| - |F_{\text{calc}}||) / \sum |F_{\text{obs}}|$, where $|F_{\text{obs}}|$ is the observed and $|F_{\text{calc}}|$ the calculated structure amplitude of a reflection after back-transformation of the averaged electron density.

‡ The r.m.s. deviation in bond length and bond angles from ideal values.

§ The r.m.s. deviation between non-crystallographic symmetry (NCS) related atom positions.

|| The r.m.s. deviation in temperature factors between bonded and non-crystallographic symmetry related atoms.

R_{free} is the same as R_{crist} calculated with a test set comprising 5% of the whole data set that was not used in the refinement.

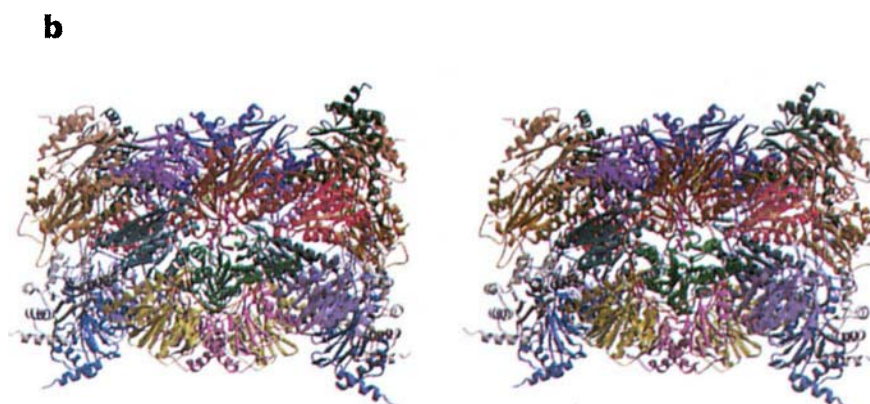
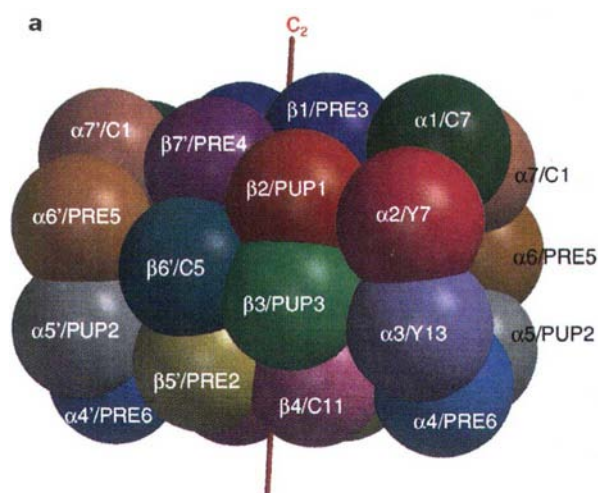


Figure 2 a, b, Topology of the 28 subunits of the yeast 20S proteasome drawn as a, spheres, and b, in ribbon representation (prepared with MOLSCRIPT, RASTER3D⁵⁴).

from their D7 symmetric ancestor complex, mirrored in the *T. acidophilum* proteasome, but despite their conserved architecture are made subunit-specific by their unique amino-acid sequences.

There are a host of additional contacts that are absent in *T. acidophilum* and which are contributed by sequences and sequence insertions characteristic for yeast and eukaryotes in general. Within the α -rings, intimate α -cis contacts are made by the intertwined N-terminal segments of subunits $\alpha 1/C7$, $\alpha 2/Y7$, $\alpha 3/Y13$ and $\alpha 7/C1$ in the centre of the heptameric rings. Tyr 8 conserved in all subunits plays a central role. Within the β -rings, a very specific contact exists between $\beta 2/PUP1$ and $\beta 3/PUP3$ mediated by the long C-terminal arm of PUP1, which embraces PUP3, touching the next-nearest neighbour $\beta 4/C11$. β -trans- α contacts are by the helix H1-loop-helix H2 motifs which interdigitate with the same motifs of two adjacent α -subunits. This basic contact geometry was observed in the *T. acidophilum* structure (see Fig. 4a in ref. 14), but the insertion at residue 66 of $\beta 7/PRE4$ favours its association with $\alpha 6/PRE5$ and $\alpha 7/C1$. Similarly, the long insertion segment in $\alpha 2/Y7$ between strands S9 and S10 binds to $\beta 2/PUP1$ and couples this pair. Specific β -trans- β interactions are formed by the $\beta 7/PRE4$ C-terminal arm which intercalates between $\beta 1'/PRE3$ and $\beta 2'/PUP1$. The C-terminal segment of $\beta 5/PRE2$ interacts with $\beta 3'/PUP3$ and $\beta 4'/C11$ in a similar manner. The long insertion of $\beta 6/C5$ at residue 145 contacts subunit $\beta 3'/PUP3$ and the C-terminal arm of $\beta 2'/PUP1$ (Fig. 3). Very specific β -trans- β interactions are mediated by magnesium ions: magnesium Y8 links the main-chain carboxylate of Asp 193 $\beta 6/C5$ with the loop 162 to 167 of $\beta 2'/PUP1$. In the same way, magnesium Y9 links $\beta 3/PUP3$ via Asp 193 to $\beta 5'/PRE2$. The aspartate residues are completely buried and their side chains involved in charge interactions with Arg 19 $\beta 2'/PUP1$ and Arg 19 $\beta 5'/PRE2$, respectively, strengthening the β -trans- β contacts further. The same C-terminal carboxylate groups are ligands to other magnesium ions located in loops 165 of $\beta 6/C5$ (magnesium W4) and $\beta 3/PUP3$ (magnesium W6), respectively, which may play a role in subunit stabilization. The β -type subunits $\beta 1/PRE3$ and $\beta 4/C11$ lie at the single molecular diad of the yeast proteasome and mimic the dominant β -trans- β contact at residues 133–137 of helix H3 of the *T. acidophilum* proteasome where all β -subunits have diad-related neighbours. This region shows a pronounced sequence similarity to *T. acidophilum*.

The N-terminal threonine site

A catalytic system formed by Thr 1, Glu 17 and Lys 33 has been defined in the *T. acidophilum* protease by structural and mutational studies^{14,15}. Close to Thr 1 are residues Ser 129, Ser 169 and Asp 166, which seem to be required for the structural integrity of the site but may also be involved in catalysis. Asp 166 is essential in the *T. acidophilum* proteasome, as shown by mutagenesis²⁴. These residues are invariant in the active subunits. In addition, we find a fully occupied solvent molecule, NUK, in all three subunits close to

Thr 1 O γ and N, Ser 129 O γ and N, and Gly 47 N, as shown for subunit $\beta 5/PRE2$ in the lactystin complex (Fig. 4a). This had been missed in the *T. acidophilum* electron density map at lower resolution, but was seen in penicillin acylase, a member of the Ntn-hydrolase family³¹. Thr 1 N is hydrogen-bonded to L168 O and Ser 169 O γ and to Ser 129 O γ , and Thr 1 O γ is hydrogen bonded to Lys 33 N ζ . Asp 17 is fully hydrogen-bonded through O $\delta 1$ to Arg 19 N and Gly 170 N and through O $\delta 2$ to Thr/Ser 2 N and Lys 33 N ζ . Similarly, Lys 33 N ζ makes three hydrogen bonds with Asp 17 O $\delta 2$, Arg 19 O and Thr 1 O γ . The pK_a and status of protonation of the ionizable groups is unknown, but the pattern of hydrogen bonds suggests that both Asp 17 and Lys 33 are charged. The positive charge on Lys 33 may shift the intrinsic pK_a of NUK and of the Thr 1 amino and hydroxyl groups, enhancing their nucleophilicity. Thr 1 N is most probably the proton acceptor when

Table 2 Nomenclature of proteasome subunits

Internal code	New systematic name	Traditional name (<i>S. cerevisiae</i>)	<i>H. sapiens</i>
G	$\alpha 1_{sc}$	C7/PRS2	iota
A	$\alpha 2_{sc}$	Y7	C3
B	$\alpha 3_{sc}$	Y13	C9
C	$\alpha 4_{sc}$	PRE6	C6
D	$\alpha 5_{sc}$	PUP2	zeta
E	$\alpha 6_{sc}$	PRE5	C2
F	$\alpha 7_{sc}$	C1/PRS1	C8
N	$\beta 1_{sc}$	PRE3	Y/delta
	$\beta 1i_{hs}$	—	LMP2
H	$\beta 2_{sc}$	PUP1	Z
	$\beta 2i_{hs}$	—	MECL1
I	$\beta 3_{sc}$	PUP3	C10
J	$\beta 4_{sc}$	C11/PRE1	C7
K	$\beta 5_{sc}$	PRE2	X/MB1
	$\beta 5i_{hs}$	—	LMP7
L	$\beta 6_{sc}$	C5/PRS3	C5
M	$\beta 7_{sc}$	PRE4	N3/beta

Adjacent subunits are numbered sequentially within each ring starting from the subunit closest to the particle diad. α and β subunits with equal numbers are adjacent. The equivalent rings and subunits related by the local diad of the particle are distinguished by the prime symbol. Thus at the interface of the β and β' rings subunit $\beta 1$ is packed against $\beta 1'$, $\beta 2$ to $\beta 7'$, $\beta 3$ to $\beta 6'$, and so on.

The exchangeable subunits in the 'immuno' proteasome are marked with an 'i'.
_{sc}, *Saccharomyces cerevisiae*; _{hs}, *homo sapiens*.

Table 3 Main-chain regions of proteasomal subunits for which electron density is defined

$\alpha 1/C7$: Gly6–Asp240; $\alpha 2/Y7$: Thr5–Leu236; $\alpha 3/Y13$: Gly4–Gly237; $\alpha 4/PRE6$: Tyr8–Gln244; $\alpha 5/PUP2$: Arg10–Glu243 (7 residues of the insertion are not defined: Gly126–Arg126); $\alpha 6/PRE5$: Phe4 Ile233; $\alpha 7/C1$: Gly5–Asn241; $\beta 1/PRE3$: Thr1–Leu196; $\beta 2/PUP1$: Thr1–Cys221; $\beta 3/PUP3$: Ser –8–Asp193; $\beta 4/C11$: Met –1–Gln192; $\beta 5/PRE2$: Thr1–Gly211; $\beta 6/C5$: Gln –9–Asp193; $\beta 7/PRE4$: Thr –8–Ile211.

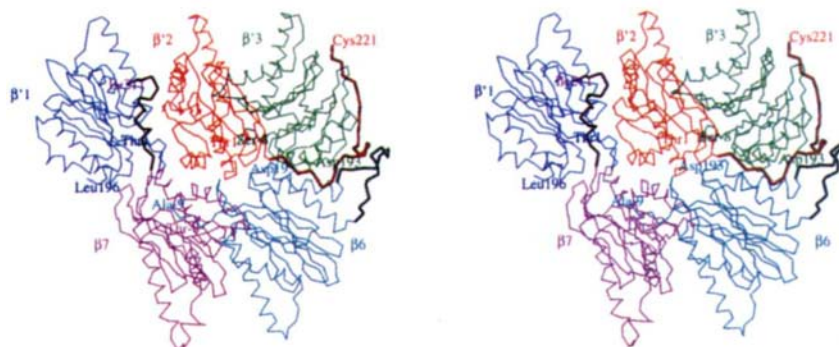


Figure 3 C α chain tracings of subunits $\beta 7/PRE4$, $\beta 6/C5$, $\beta 1'/PRE3$, $\beta 2'/PUP1$ and $\beta 3'/PUP3$, highlighting their β -cis and β -trans- β interactions through contacts of insertion segments.

Thr 1 O γ adds to an electrophilic centre. This is well illustrated by the structure of the complex with lactacystin, which acylates Thr 1 O γ as a result of β -lactone ring opening. Thr 1 N is ideally positioned to serve as a proton shuttle from Thr 1 O γ to the lactacystin O6 for this process. An analogous sequence of chemical steps is proposed for hydrolysis of the C-terminal fluorophore of fluorogenic substrates, whereby proton transfer is to the leaving group's amido nitrogen. The acyl enzyme generated may be deacylated by water NUK (Fig. 5d–f). Alternatively, or in parallel, there might be direct attack of the peptide bond by NUK, bypassing intermediate e.

Inhibitor binding and specificity

β 1/PRE3, β 2/PUP1, and β 5/PRE2 have acetyl-Leu-Leu-norleucinal (calpain inhibitor I) covalently bound to Thr 1 O γ , probably as a hemiacetal. It adopts a β -conformation and fills a gap between

strands, including residues 20, 21 (assigned to the loop L in 14) and 47, respectively, to which it is hydrogen-bonded, generating an antiparallel β -sheet structure. The norleucine side chain projects into a pocket (the S1 pocket) that opens sideways towards a tunnel leading to the particle surface, the leucine side chain at P2 is not in contact with protein, and the leucine side chain at P3 is in contact with the adjacent β -type subunit. The S1 ('specificity') pocket is mainly shaped by residues Thr 20, Thr 31, Thr 35, Arg 45, Ala 49 and Gln 53 in β 1/PRE3 (Fig. 6a); Ser 20, Cys 31, His 35, Gly 45, Ala 49 and Glu 53 in β 2/PUP1 (Fig. 6b); and Ala 20, Val 31, Ile 35, Met 45, Ala 49 and Gln 53 in β 5/PRE2 (Fig. 6c). Residue 45 forms the bottom of the pocket and appears largely to determine its character. Adjacent subunits in the β -rings contribute to the S1 pockets further back and modulate their character: β 2/PUP1 in the case of β 1/PRE3, with His 114, His 116, Ser 118, Asp 120, β 3/PUP3 in

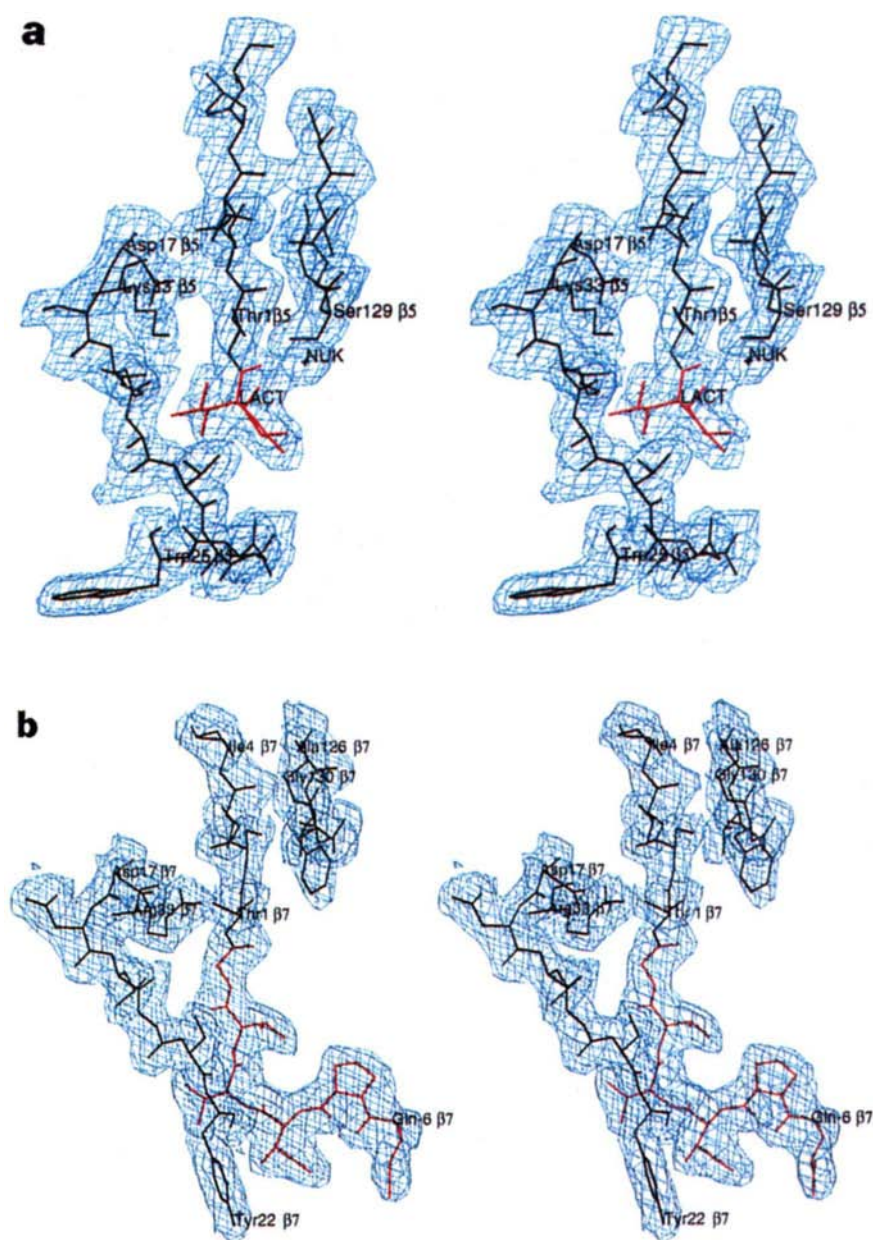


Figure 4 a, b. Electron density maps (contoured from 1σ) in similar orientations around Thr 1 with $2F_o - F_c$ coefficients after two-fold averaging. The red model parts have been omitted for phasing. **a**, β 5/PRE2 with the covalently bound

lactacystin (LACT) and the water molecule NUK; **b**, β 7/PRE4 with part of its propeptide.

the case of $\beta 2/\text{PUP1}$, with residues Asp 114, Asp 120 and Cys 118 and $\beta 6/\text{C5}$ in the case of $\beta 5/\text{PRE2}$ with Ser 118, Asp 114, Glu 120 and Glu 122.

Lactacystin is covalently bound to $\beta 5/\text{PRE2}$, in agreement with the observed chemical modification of subunit X of the mammalian proteasome³², the homologue of PRE2. Its dimethyl side chain at C10 projects into S1 like a valine of leucine side chain, but not so deeply as the norleucine side chain of the calpain inhibitor. Lactacystin displays a host of hydrogen bonds with protein main-chain atoms Lact N–Gly 47O, LactO4–Gly47N, LactO9–Thr 21N, LactO6–Thr1N. As these hydrogen-bonding interactions could also

be made in $\beta 1/\text{PRE3}$ and $\beta 2/\text{PUP1}$, which do not form covalent complexes with lactacystin, the side group that binds in the hydrophobic S1 pocket of $\beta 5/\text{PRE2}$ seems to direct covalent bond formation and stabilization and is an obvious target for inhibitor development. Met 45 of $\beta 5/\text{PRE2}$ is in contact with the branched side chain of lactacystin in that complex. In the calpain inhibitor complex, the norleucine side chain pushes the methionine side chain up by 2.7 Å towards Ile 35, which rotates out of the way. These concerted movements make the S1 pocket more spacious. Its acidic side walls, contributed by subunit $\beta 6/\text{C5}$, may also allow binding of basic residues, in particular arginine. This is consistent with the

Figure 5 Scheme of the proposed chemical steps of autolysis and substrate hydrolysis. Generation of a processing intermediate by hydrolysis at the 'acidic' β -annulus (a). Generation of the fully processed active subunit by an acyl-enzyme (b) and its hydrolysis (c). Michaelis complex of a substrate polypeptide (d). Cleavage at the β -annulus and formation of the acyl-enzyme associated with peptide cleavage (e). Acyl-enzyme hydrolysis and liberation of the octapeptide product (f).

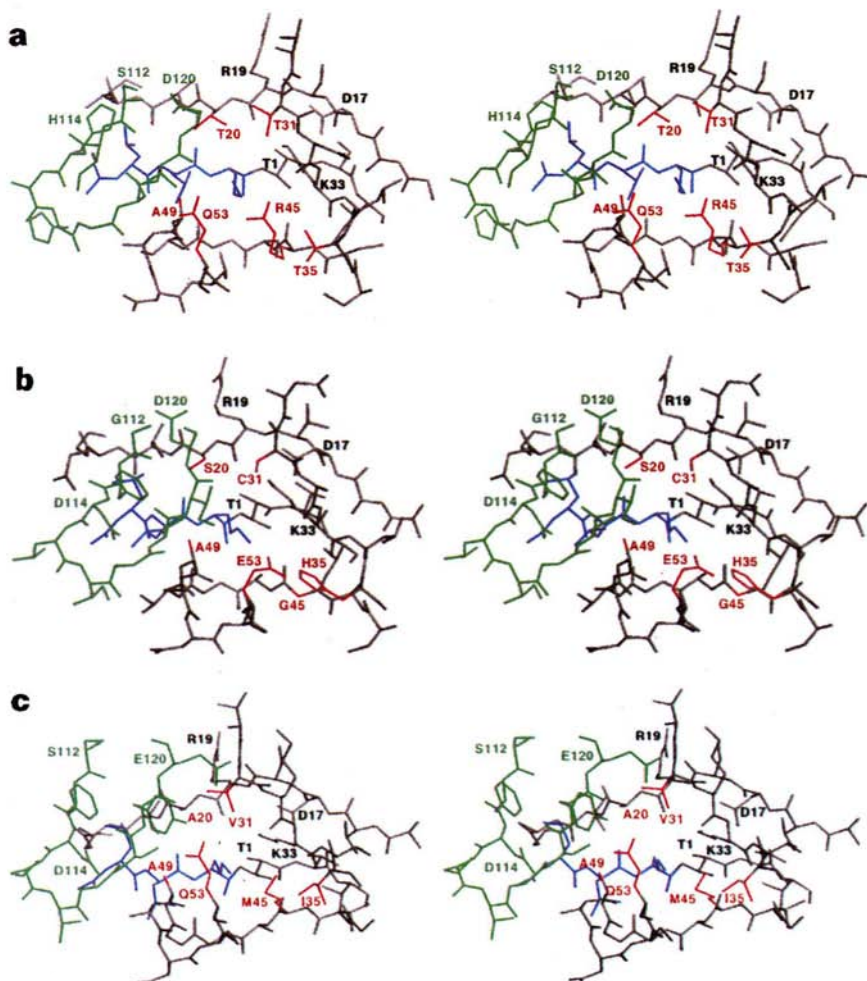
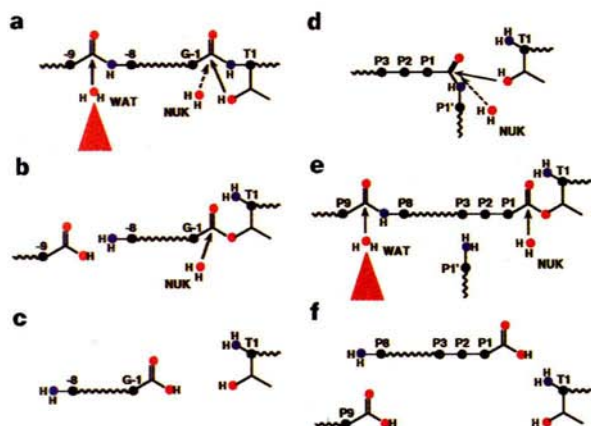


Figure 6 a–c, Calpain inhibitor binding and S1 pockets. **a**, $\beta 1/\text{PRE3}$ in grey, with residues contacting P1 in red, $\beta 2/\text{PUP1}$ in green and the inhibitor in blue; **b**, $\beta 2/\text{PUP1}$, $\beta 3/\text{PUP3}$ and **c**, $\beta 5/\text{PRE2}$, $\beta 6/\text{C5}$ with analogous colour codes (drawn with MAIN (ref. 50)).

observation that lactacystin inhibits both the chymotrypsin-like and trypsin-like activity against chromogenic substrates³². The chymotrypsin-like activity is also reduced in proteasomes carrying a mutant form of $\beta 5/\text{PRE2}$ that cannot be processed from its proform²⁵ and those with a mutation in $\beta 5/\text{PRE2}$ (ref. 33), where a substitution of Ala 49 by Val in the S1 pocket restricts its size. $\beta 1/\text{PRE3}$, with Arg 45 at the base of S1, is well suited for glutamate P1 residues and probably the subunit associated with the PGPH activity of the proteasome, as indicated by mutational analysis¹³. But the norleucine side chain also occupies this pocket in the calpain inhibitor complex. We observe a high additional density peak associated with the guanidinium side chain, which we interpret as being due to a chloride ion compensating for the unbalanced positive charge. $\beta 2/\text{PUP1}$ has a glycine at residue 45 and consequently a spacious S1 pocket confined at its bottom by His 35.

We conclude that $\beta 5/\text{PRE2}$ carried both chymotrypsin-like and trypsin-like specificity and that $\beta 1/\text{PRE3}$ carries the PGPH activity, but both pockets are adaptable in size (PRE2) and polarity (PRE3). $\beta 2/\text{PUP1}$ is suited for very large P1 residues of neutral or acidic or basic character, depending on whether His 35 or the acidic $\beta 3/\text{PUP3}$ side wall provide the countercharge.

Mutational analysis has shown that substitutions in $\beta 4/\text{C11}$ and $\beta 7/\text{PRE4}$ affect the chymotrypsin-like and PGPH activity, respectively^{13,33,34}. These subunits are inactive but situated adjacent to the $\beta 5/\text{PRE2}$ and $\beta 1/\text{PRE3}$ subunits from both rings, respectively (Fig. 3). The exchange in $\beta 4/\text{C11}$ of the internal Ser 136 by the bulky phenylalanine residue disrupts the β -*trans*- β contact at helix H3 between $\beta 4/\text{C11}$ and $\beta 5/\text{PRE2}$ and may distort the adjacent Thr 1 site, as presumably does the deletion of 15 C-terminal residues from $\beta 7/\text{PRE4}$, which form extensive contacts with $\beta 1/\text{PRE3}$ (Fig. 3).

Propeptides and endopeptidase activity

Five β -type subunits of the yeast proteasome are synthesized with propeptides of up to 75 residues which are cleaved during proteasome maturation. $\beta 1/\text{PRE3}$, $\beta 2/\text{PUP1}$ and $\beta 5/\text{PRE2}$ autolyse between Gly-1–Thr1 in a process that requires the presence of Thr 1, Gly -1 and Lys 33, as shown for $\beta 5/\text{PRE2}$ and LMP2, the mammalian homologue of $\beta 1/\text{PRE3}$ (refs 25, 35). The agents in catalysis are the same as those in the chromogenic substrate activity, except for the N terminus of Thr 1, which is not available in the proforms. We previously suggested that intrasubunit autolysis occurs by Thr 1 O γ acting as the nucleophile and attacking the preceding peptide bond³⁵. The crystal structure assigns a key role to water NUK. It is ideally positioned to act as a general base to abstract a proton from Thr 1 O γ and drive nucleophilic addition to the carbonyl carbon of Gly -1. We do not know the positions and orientation of the Gly-1–Thr1 peptide groups of the subunits that

have been fully processed, but they are probably similar to the generally conserved geometries of the unprocessed or intermediately processed subunits $\beta 3/\text{PUP3}$, $\beta 6/\text{C5}$ and $\beta 7/\text{PRE4}$. Gly-1 O is directed towards Lys 33 N ζ and Gly 47 N in these subunits, which represent an oxyanion hole that can accommodate the negative charge that develops upon tetrahedral adduct formation, as in the serine proteases. Collapse to the ester may follow after proton transfer from water NUK to Thr 1 N and cleavage of the peptide bond. Ser 129 O γ and Ser 169 O γ are nearby to assist in this reaction. Both hydroxyl groups are hydrogen-bonded to Asp 166, which is invariant in the active subunits. NUK is probably also involved in ester hydrolysis as the attacking nucleophile that would finally be incorporated into product (Fig. 5a–c). Glycine at position -1 seems to be essential, because a side chain would interfere with the protein backbone at residue 168 and may induce a configuration unsuitable for autolysis.

With the generation of the N terminus at Thr 1, the subunits become active. Autolysis at residue 1 does not occur if the catalytic site is not intact, as in subunits $\beta 3/\text{PUP3}$, $\beta 4/\text{C11}$ and $\beta 6/\text{C5}$, which lack Thr 1, and in $\beta 7/\text{PRE4}$ which has Lys 33 replaced by Arg, and in mutants of LMP2, the mammalian homologue of $\beta 1/\text{PRE3}$ (ref. 35), and of $\beta 5/\text{PRE2}$ (ref. 25). $\beta 7/\text{PRE4}$ has both Gly -1 and Thr 1, but not in the same conformation as in the active subunits, because the Thr 1 side chain is pushed away by the bulkier Arg 33 (Fig. 4b). In addition to the chymotrypsin-like activity³³, processing²⁵ is abolished by a mutation in the adjacent subunit $\beta 4/\text{C11}$, Ser 136 to Phe. These defects in catalytic activity and in processing are testimony to the structural lability of the Thr 1 site, which may be distorted by mutations of neighbouring residues in the same or in adjacent subunits. Conversely, an inactive mutant may regain activity in the environment of active subunits, like *T. acidophilum* species that are defective in processing but which can be processed when coexpressed with wild-type protein²⁴.

$\beta 6/\text{C5}$ and $\beta 7/\text{PRE4}$ are both processed to intermediates and the location of the cleavage sites should lead to the active sites responsible for this endopeptidase activity, assuming that there is no rearrangement after cleavage. A gross rearrangement is unlikely as the propeptides are held tightly to the protein walls, proPRE4 to subunits $\beta 7/\text{PRE4}$ and $\beta 1/\text{PRE3}$ and proC5 to subunits $\beta 6/\text{C5}$ and $\beta 7/\text{PRE4}$. Residues Thr -8 and Gln -9 lie very close to one another at the sharp-edged inner annulus of the β -rings, which is formed by the C termini of helices H2 of the seven subunits. Specifically, Thr -8 and Gln -9 of $\beta 7/\text{PRE4}$ and $\beta 6/\text{C5}$, respectively, are close to helices H2 of subunits $\beta 1/\text{PRE3}$ and $\beta 7/\text{PRE4}$ (Fig. 7). The putative processing intermediates of the $\beta 1/\text{PRE3}$, $\beta 2/\text{PUP1}$ and $\beta 5/\text{PRE2}$ subunits would also lie on the inner annulus of the β -ring in a conformation similar to that found in $\beta 7/\text{PRE4}$ and $\beta 6/\text{C5}$. Intermediate processing

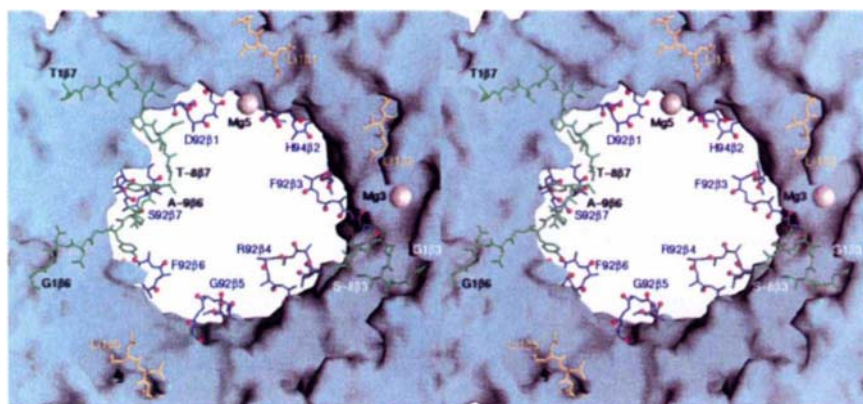


Figure 7 The lower half of the β - β chamber. The main chain with red spheres for the carbonyl oxygens is traced for the C-terminal parts of helices H2 of the seven β -type subunits defining the β -annulus. The intermediately processed and the unprocessed propeptides of subunits $\beta 6/\text{C5}$, $\beta 7/\text{PRE4}$, $\beta 3/\text{PUP1}$ and $\beta 4/\text{C11}$

(green) and the calpain inhibitors (yellow) bound to $\beta 1/\text{PRE3}$, $\beta 2/\text{PUP1}$ and $\beta 5/\text{PRE2}$ are shown. Two magnesium ions located close to the β -annulus are shown as silver spheres (drawn with GRASP⁵⁵).

is observed for mutant LMP2 with an impaired catalytic apparatus³⁵ and probably also for $\beta 5/\text{PRE}2$, as suggested from the molecular masses of intermediates after pulse-chase experiments (Figs 2b, 3b in ref. 25). This is remarkable because $\beta 5/\text{PRE}2$ has no adjacent active neighbour in the mature 20S proteasome.

Proteasomes have a very broad, but length-restricted, sequence specificity¹⁸ which generates peptides of ~ 8 residues. Based on the *T. acidophilum* 20S proteasome structure, we have suggested that the distance of 28 Å between the Thr1 active sites of neighbouring subunits determines the length of the peptide product generated by sequential processive proteolysis^{9,10,14,36}. The fact that processing intermediates exist suggests that, in addition to the N-terminal threonine residue being the site of autolysis and cleavage of short chromogenic substrates, there is a second hydrolytic site, which we tentatively localize at the inner β -ring annulus. We propose that a substrate polypeptide docks at the Thr1 site with a hydrogen-bonding pattern similar to the tripeptide aldehyde inhibitor binding. Acyl-enzyme formation may require the cleavage of a preceding peptide bond as in the autolysis reaction, and proteolysis may therefore start from an internal site. Steric constraints would require a sharply bent conformation at the scissile peptide bond in this case, and the reaction would proceed as shown in Fig. 5d–f. Residues –9 or –8 would lie at the inner annulus when bound like $\beta 6/\text{C}5$ and $\beta 7/\text{PRE}4$ propeptides. We suggest that water molecules activated by the 3 to 4 main-chain carbonyl oxygens at the helix ends of any of the seven subunits (Fig. 6) act as the general base and acid leading to hydrolysis, rather like the aspartate proteases, in which carboxylic acid groups activate a water molecule. Eighteen magnesium-binding sites were identified in the proteasome molecule, of which 12 are located on the inner walls of the β – β chamber. These testify to the acidic nature of this compartment. There is no obvious peptide-binding pocket at these sites, so cleavage may not be specific—agreeing with the apparent lack of specificity against longer peptide substrates^{18–21}.

The propeptides are essential for the assembly of eukaryotic proteasomes^{23,25}, which may be because of their participation in intersubunit contacts, stabilization of subunit structures, or both. The structures of the processing intermediates of $\beta 6/\text{C}5$ and $\beta 7/\text{PRE}4$ and the unprocessed $\beta 3/\text{PUP}3$ propeptides provide a clue and suggest that both effects do occur, as the propeptides are tightly attached to the rest of the protein and interact with other subunits, $\text{pro}\beta 7/\text{PRE}4$ with $\beta 1/\text{PRE}3$ at residues 92 and 115 and $\text{pro}\beta 6/\text{C}5$ with $\beta 7/\text{PRE}4$ at residues 91 and 116.

Entry into and exit from the proteasome

The hydrolytic activity of the proteasome is associated with Thr1 and, hypothetically, the β -annuli inside the β -cavity that defines the hydrolytic chamber. Substrate has to enter the particle and product must be released. In the *T. acidophilum* proteasome, two entry ports

with a diameter of ~ 13 Å at the ends of the cylindrical particles are open, constricted by an annulus of turn-forming segments Tyr126–Gly–Gly–Val of the seven identical α -subunits. The N-terminal residues 1 to 12 are disordered in that protein. In contrast, the hydrolytic chamber of the yeast 20S proteasome is quite inaccessible. The N-termini of subunits $\alpha 1/\text{C}7$, $\alpha 2/\text{Y}7$, $\alpha 3/\text{Y}13$, $\alpha 6/\text{PRE}5$ and $\alpha 7/\text{C}1$ project into the opening and fill it completely in several layers with tightly interdigitating side chains (Fig. 8). There is no access into the particle interior from the cylinder ends without substantial rearrangement. There are some narrow side windows, particularly at the interface between the α - and β -rings, which seem to be more permeable than in the *T. acidophilum* proteasome, because smaller side chains participate. These openings are mainly in between the teeth-like helix H1–turn–helix H2 motifs of the α – β interface (Fig. 4a in ref. 14) and lead to the active site threonines. They are coated with polar and charged residue side chains, which may move to make openings of ~ 10 Å diameter and possibly allow passage of unfolded, extended polypeptide chains. The 19S particle, which confers ATP and ubiquitin dependence on proteolysis by the proteasome, is attached to the particle poles to form the 26S proteasome. This association results in a strong activation of peptide hydrolysis³⁷. Similarly, the proteasome regulator PA28 is bound to α -type subunits^{38,39}. It accelerates peptide digestion and improves antigen processing¹². Both regulatory factors may open entry ports in a controlled manner *in vivo*.

MHC class I peptide generation

The 20S proteasome generates peptide products with a narrow length distribution centred around octa- and nonapeptides, sizes suitable for binding to MHC class I molecules⁴⁰. Peptides generated by 20S proteasomes from intact proteins are presented by MHC class I molecules *in vitro*^{19,41}. *In vivo*, proteasome inhibitors suppress MHC class I presentation of protein antigens⁴². Further, the number of cell-surface MHC class I molecules is controlled by the inducible proteasomal subunits $\beta 5i/\text{LMP}7$ and $\beta 1i/\text{LMP}2$, which replace constitutively expressed subunits, as shown in mice with a targeted deletion of the gene encoding LMP7 (ref. 29) and in a T-cell lymphoma in which $\beta 1i/\text{LMP}2$ is underexpressed⁴³.

MHC class I peptides usually have basic or hydrophobic C-terminal residues⁴⁴ and the LMP2/7 substitution may alter the distribution of peptides so that more MHC class I-preferred peptides are generated. LMP2 replaces Y, the human homologue of $\beta 1/\text{PRE}3$ and LMP7 replaces X, the homologue of $\beta 5/\text{PRE}2$. All subfamily members share extensive sequence identity, but $\beta 1i/\text{LMP}2$ has two Thr 31 $\rightarrow$ Phe and Arg 45 $\rightarrow$ Leu replacements with respect to $\beta 1/\text{PRE}3$ and Y in the S1 pocket (Fig. 6a–c). The pocket becomes apolar and more constricted so that the PGPH activity should be reduced and the chymotryptic activity enhanced, as has

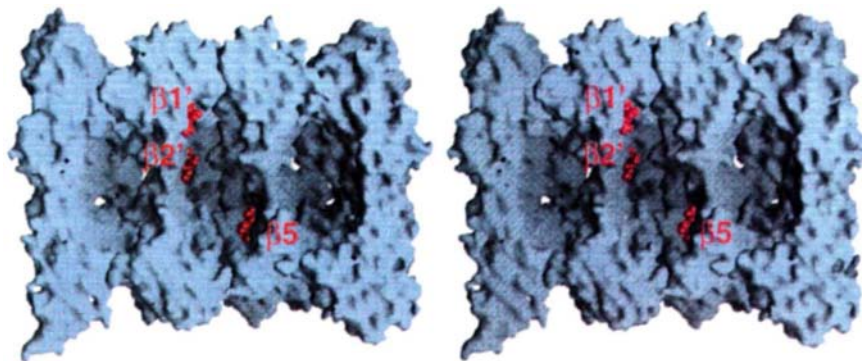


Figure 8 Surface view of the proteasome molecule clipped along the cylinder axis. Three of the six calpain inhibitor molecules bound to $\beta 1'/\text{PRE}3$, $\beta 2'/\text{PUP}1$ and $\beta 5/\text{PRE}2$ are shown as space-filling models in red. The sealed α -ports at both

ends of the particle, a few narrow side windows, and the sharply cut inner β -annuli can be seen (drawn with GRASP⁵⁵).

been observed^{27,28}. The opposite effect is seen in cell lines that lack LMP2 and LMP7 genes and in mutant mice with a disrupted LMP2 gene³⁰. Replacement of X and Z, the human homologues of $\beta 5/\text{PRE2}$ and $\beta 2/\text{PUP1}$, with LMP7 and MECL1, respectively, do not alter the S1 pockets and it is unclear from the structure why LMP7 influences the MHC class I presentation pathway *in vivo* in mutant mice²⁹ and their tissues⁴⁵.

The processing intermediates of subunits $\beta 7/\text{PRE4}$ and $\beta 6/\text{C5}$ are octa- or nonapeptide products that are not liberated from the enzyme. Both peptides have similar conformations, with a bulge dividing two parts each in an extended conformation, reminiscent of the conformation of peptides bound to MHC class I molecules. Superposition of the propeptide of $\beta 6/\text{C5}$ with a viral peptide nonamer in complex with its MHC class I receptor⁴⁶ quantifies the similarity, with r.m.s. deviations for all atoms of 2.3 Å and for C α atoms of 1.3 Å; it also indicates that preferred local conformations of the unfolded polypeptide might participate in the generation (by the proteasome) and presentation (by MHC class I molecules) of immunodominant peptide epitopes. □

Methods

Protein preparation, characterization and crystallization. Yeast cells of *S. cerevisiae* (Hefe-Mayr) were washed twice with ice-cold water and suspended in buffer A (20 mM Tris-HCl, pH 7.5, 1 mM EDTA and 1 mM Na₂S₂O₃). The ratio of cells to buffer was 2/3 (w/w). Cells were broken for 10 min in a disintegrator (Biomatik) with glass beads (diameter: 0.5 mm; ratio of beads to cell suspension: 3/2 (v/v)). Cell breakage was monitored microscopically. After filtration, the crude extract was centrifuged for 45 min at 10,000g in a Sorvall RC 2B centrifuge. The supernatant was recentrifuged for 60 min at 134,000g in a Ti-55.2 rotor (Beckman). Lipids from the top layer were removed and the remaining yellow solutions combined. Protein concentrations were ~50 mg ml⁻¹. Immediately after centrifugation, the extract was applied to a Q-Sepharose column (5 × 20 cm) equilibrated with 280 mM NaCl in buffer A. The column was washed with 280 mM NaCl in buffer A and proteins were eluted with 2 litres of buffer solution using a gradient of 280–800 mM NaCl. The flow rate was 120 ml h⁻¹ and 12-ml fractions were collected. The proteasome eluted at 400–450 mM NaCl. Chymotrypsin-like (CL), peptidylglutamyl peptide-hydrolase (PGPH) and trypsin-like (TL) activity was measured in all fractions. Pooled fractions were diluted threefold with water and applied to a hydroxy-apatite column (3 × 10 cm) equilibrated with 60 mM potassium phosphate, pH 7.5. The column was washed with 60 mM potassium phosphate, pH 7.5, and eluted with 1 litre of buffer solution with a gradient of 60–300 mM of potassium phosphate. The flow rate was 60 ml h⁻¹ and 12-ml fractions were collected. The proteasome eluted at 120 mM phosphate. CL, PGPH and TL activity was measured in all fractions and active fractions were combined. The pooled fractions were concentrated 20-fold by ultrafiltration using an AMICON YM30 membrane and the concentrate was applied to a Superose-6 column (1 × 30 cm) equilibrated with buffer A. Elution was at a flow rate of 18 ml h⁻¹ in buffer A. The proteasome eluted at 37 min. All preparative steps except FPLC were carried out at 4 °C. The chromogenic peptide substrates were dissolved in dimethylsulphoxide at 1 mM. Proteolytic activity against these substrates was determined according to ref. 47. Chromogenic peptide substrates were obtained from Bachem (Bubendorf, Switzerland), Q-Sepharose and hydroxy-apatite from Sigma and Biorad, respectively, FPLC equipment and Superose-6 columns from Pharmacia. All other chemicals (highest purity) were from Merck.

Crystals were grown in hanging drops at 24 °C. The protein concentration used for crystallization was 40 mg ml⁻¹ in 10 mM Tris-HCl (pH 7.5) and 1 mM EDTA. The drops contained 4 μ l protein and 2 μ l reservoir solution, with 40 mM magnesium acetate, 0.1 M morpholino-ethane-sulphonic acid (pH 6.5) and 12% 2,4-methylpentanediol. Crystals of lactacystin (LACT; kindly provided by S. L. Schreiber and E. J. Corey, Harvard) complex were prepared by soaking in a 1 mM solution for 6 h and of the acetyl-Leu-Leu-Norleucinal (CAL, or calpain inhibitor I; Boehringer-Mannheim) complex by soaking in a 5 mM solution for 6 h.

Crystallographic methods and data. Data sets were collected at BW6 beamline at the DESY/Hamburg with synchrotron radiation of $\lambda = 1.1$ Å.

Crystals were soaked in a cryoprotectant buffer (30% MPD, 28 mM magnesium acetate, 0.1 M morpholino-ethane-sulphonic acid, pH 6.9) and frozen in a stream of cold nitrogen gas at 90 K (Oxford Cryosystems). Diffraction data were collected with a 300-mm Mar research imaging plate at a distance of 275 mm (LACT) or 280 mm (CAL). X-ray intensities were evaluated by using the MOSFLM program package (version 5.3) and data reduction was performed with CCP4 (ref. 48). The Patterson search calculations were done with AMoRe⁴⁹. Averaging of electron density maps in real space was carried out with MAIN⁵⁰. Anisotropy of diffraction was corrected by comparing observed structure-factor amplitudes with those calculated from a model with isotropic temperature factors using X-PLOR⁵¹. Model building was done on an ESV-30 Graphic system workstation (Evans & Sutherland) using FRODO⁵². Crystallographic refinement and electron density map calculations were done with X-PLOR⁵¹ using tight energy (with the parameters in ref. 53) and 2-fold non-crystallographic symmetry restraints.

Amino-acid sequences. The amino-acid sequences of the mature yeast proteasome subunits and of the human LMP2, LMP7, MECL1 were taken from SWISSPROT (details from the authors on request). The amino-acid numbering and the secondary structure follow the *T. acidophilum* structure¹⁴. We propose structure-based names for the proteasomal subunits in Table 2 and use them together with the traditional nomenclature.

Received 31 December 1996; accepted 13 February 1997.

- Emori, Y. *et al.* *Mol. Cell. Biol.* **11**, 344–353 (1991).
- Rivett, A. J. *Biochem. J.* **291**, 1–10 (1993).
- Peters, J.-M. *Trends Biochem. Sci.* **19**, 377–382 (1994).
- Goldberg, A. L., Stein, R. & Adams, J. *Chem. Biol.* **2**, 503–508 (1995).
- Hochstrasser, M. *Curr. Opin. Cell Biol.* **7**, 215–223 (1995).
- Rubin, D. M. & Finley, D. *Curr. Biol.* **5**, 854–858 (1995).
- Lord, M. *Curr. Biol.* **6**, 1067–1069 (1996).
- Coux, O., Tanaka, K. & Goldberg, A. L. *Annu. Rev. Biochem.* **65**, 801–847 (1996).
- Stock, D. *et al.* *Curr. Opin. Biotech.* **7**, 376–385 (1996).
- Stock, D., Ditzel, L., Baumeister, W., Huber, R. & Lowe, J. *Cold Spring Harbor Symp. Quant. Biol.* **LX**, 525–532 (1995).
- Chen, P. & Hochstrasser, M. *EMBO J.* **14**, 2620–2630 (1995).
- Groettrup, M., Soza, A., Kuckelkorn, U. & Kloetzel, P. M. *Immunol. Today* **17**, 429–435 (1996).
- Hilt, W. & Wolf, D. H. *Trends Biochem. Sci.* **21**, 96–102 (1996).
- Lowe, J. *et al.* *Science* **268**, 533–539 (1995).
- Seemüller, E. *et al.* *Science* **268**, 579–582 (1995).
- Brannigan, J. A. *et al.* *Nature* **378**, 416–419 (1995).
- Heinemeyer, W., Tröndle, N., Albrecht, G. & Wolf, D. H. *Biochemistry* **33**, 12229–12237 (1994).
- Ehring, B., Meyer, T. H., Eckerskorn, C., Lottspeich, F. & Tampé, R. *Eur. J. Biochem.* **235**, 404–415 (1996).
- Dick, L. R. *et al.* *Immunology* **152**, 3884–3894 (1994).
- Wenzel, T., Eckerskorn, C., Lottspeich, F. & Baumeister, W. *FEBS Lett.* **349**, 205–209 (1994).
- Kuckelkorn, U. *et al.* *Eur. J. Immunol.* **25**, 2605–2611 (1995).
- Frentzel, S., Pesold-Hurt, B., Seelig, A. & Kloetzel, P.-M. *J. Mol. Biol.* **236**, 975–981 (1994).
- Cerundolo, V., Kelly, A., Elliott, T., Townsley, J. & Townsend, A. *Eur. J. Immunol.* **25**, 554–562 (1995).
- Seemüller, E., Lupas, A. & Baumeister, W. *Nature* **382**, 468–470 (1996).
- Chen, P. & Hochstrasser, M. *Cell* **86**, 961–972 (1996).
- Kopp, F. *et al.* *J. Mol. Biol.* **248**, 264–272 (1995).
- Gaczynska, M., Rock, K. L. & Goldberg, A. L. *Nature* **365**, 264–267 (1993).
- Driscoll, J., Brown, M. G., Finley, D. & Monaco, J. J. *Nature* **365**, 262–264 (1993).
- Fehling, H. J. *et al.* *Science* **265**, 1234–1237 (1994).
- Van Kaer, L. *et al.* *Immunity* **1**, 533–541 (1994).
- Duggleby, H. J. *et al.* *Nature* **373**, 264–268 (1995).
- Fenteany, G. *et al.* *Science* **268**, 726–730 (1995).
- Heinemeyer, W., Grubler, A., Möhrle, V., Mahé, Y. & Wolf, D. H. *J. Biol. Chem.* **268**, 5115–5120 (1993).
- Hilt, W., Enenkel, C., Grubler, A., Singer, T. & Wolf, D. H. *J. Biol. Chem.* **268**, 3479–3486 (1993).
- Schmidtko, G. *et al.* *EMBO J.* **15**, 6887–6898 (1996).
- Dick, L. R., Moomaw, R., DeMartino, G. N. & Slaughter, C. A. *Biochemistry* **30**, 2725–2734 (1991).
- Hoffman, L. & Rechsteiner, M. *J. Biol. Chem.* **269**, 16890–16895 (1994).
- Kania, M. A., DeMartino, G. N., Baumeister, W. & Goldberg, A. L. *Eur. J. Biochem.* **236**, 510–516 (1996).
- Gray, C. W., Slaughter, C. A. & DeMartino, G. N. *J. Mol. Biol.* **236**, 7–15 (1994).
- York, I. A. & Rock, K. L. *Annu. Rev. Immunol.* **14**, 369–396 (1996).
- Niedermann, G. *et al.* *Proc. Natl Acad. Sci. USA* **93**, 8572–8577 (1996).
- Rock, K. L. *et al.* *Cell* **78**, 761–771 (1994).
- Sibille, C. *et al.* *Curr. Biol.* **5**, 923–930 (1995).
- v. Engelhard, H. *Curr. Opin. Immunol.* **6**, 13–23 (1994).
- Stohwasser, R., Kuckelkorn, U., Kraft, R., Kostka, S. & Kloetzel, P.-M. *FEBS Lett.* **383**, 109–113 (1996).
- Madden, D. R., Gorga, J. C., Strominger, J. L. & Wiley, D. C. *Nature* **353**, 321–325 (1991).
- Achstatter, T., Ehmann, C., Osaki, A. & Wolf, D. H. *J. Biol. Chem.* **259**, 13344–13348 (1994).
- Leslie, A. G. W. Collaborative Computational Project No. 4. *Acta Crystallogr. D* **50**, 760–763 (1994).
- Navaza, J. *Acta Crystallogr. A* **50**, 157–163 (1994).
- Turk, D. Thesis, Technische Univ., München (1992).
- Brunger, A. *X-PLOR Version 3.1. A System for X-Ray Crystallography and NMR* (1992).
- Jones, T. A. *J. Appl. Cryst.* **11**, 268–272 (1978).
- Engh, R. A. & Huber, R. *Acta Crystallogr. A* **47**, 392–400 (1991).
- Kratulis, P. J. *J. Appl. Crystallogr.* **24**, 946–950 (1991).
- Nicholls, A., Bharadwaj, R. & Honig, B. *Biophys. J.* **64**, A166 (1993).

Acknowledgements. We thank P. M. Kloetzel for sharing his extensive knowledge in the proteasome field with us, M. Schneider for help with the graphic work, and R. Engh for advice. This work was supported by the Sonderforschungsbereich 207 of the Deutsche Forschungsgemeinschaft.

Correspondence and requests for materials should be addressed to R.H. The coordinates have been deposited with the Protein Data Bank Brookhaven, NY, under accession number 1RYP.

A starburst origin of the OH-megamaser emission from the galaxy Arp220

C. J. Skinner*, H. A. Smith†, E. Sturm‡, M. J. Barlow§, R. J. Cohen|| & G. J. Stacey¶

* Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, Maryland 21218, USA

† Smithsonian Institution, Laboratory for Astrophysics, Room 3726, National Air and Space Museum, Mail Stop 321, Washington DC 20560, USA

‡ Max-Planck-Institut für Extraterrestrische Physik, Postfach 1603, 85740 Garching, Germany

§ Department of Physics & Astronomy, University College London, Gower Street, London WC1E 6BT, UK

|| Nuffield Radio Astronomy Laboratories, Jodrell Bank, Macclesfield, Cheshire SK9 8DL, UK

¶ Cornell University, Department of Astronomy, 512 Space Sciences Building, Ithaca, New York 14853, USA

Ultraluminous infrared galaxies have been known for more than a decade, but the source of their very large far-infrared luminosities remains controversial. It may reflect a quasar-like active nucleus surrounded by a torus of dense gas and dust, the latter absorbing the energetic photons from the nuclear region and re-emitting at infrared wavelengths¹, or a huge burst of massive-star formation in dense dusty clouds of molecular gas close to the nucleus², which heats the surrounding dust. A number of ultraluminous galaxies are also a source of OH-megamaser emissions (intense laser-like spectral lines at microwave frequencies), an observation that may hold important clues as to the main power source in these galaxies. A general feature of many models^{3,4} is that the masers are pumped radiatively by the absorption of infrared photons. Identifying the source of the maser pump may therefore indicate whether the ultimate energy source is a burst of star formation, or an active nucleus. Here we report the detection of a strong mid-infrared OH absorption line in the prototypical megamaser and ultraluminous galaxy⁵, Arp220. We find that the power absorbed in this line alone is sufficient to pump the megamaser emission seen at radio wavelengths. Moreover, the warm, extended nature of the pumping region is suggestive of a starburst origin for the ultraluminous infrared emissions.

Details of the mechanism generating the maser emission in astrophysical sources (megamaser galaxies and, in our own galaxy, star formation regions and winds from evolved stars) are still uncertain. For the megamaser galaxies a good correlation exists between OH maser luminosity and the square of the far-infrared luminosity, especially that measured in the IRAS 60- μm band^{6,7}, consistent with an infrared radiative pump. In fact, many OH-megamaser galaxies are ultraluminous infrared galaxies (ULIRGs: 8–1,000 μm luminosities greater than $10^{12} L_{\odot}$, where $L_{\odot}$ is the solar luminosity), and all have huge far-infrared luminosities. It has been suggested that ULIRGs may be created by the merger of two galaxies¹, and Arp220 indeed appears to be the result of such a process^{8,9}. Very-long-baseline interferometry (VLBI) radio images^{10,11} of Arp220 show that the OH-megamaser emission arises in regions only a few parsecs in size in its core, and have been used to support active galactic nucleus (AGN) models for this galaxy.

The OH radical has strong transitions at 35 μm and 53 μm , and it is suspected that the masers may be pumped radiatively via absorption in these lines, exploiting the powerful infrared emission of the host galaxy. Such models^{3,4} posit that absorption of photons via the above two transitions excites the OH radicals. There follows a radiative cascade to lower levels via other far-infrared transitions,

principally those at 79, 96, 98, 119 and 163 μm , with transitions from the 35- μm pumped level ($^2\Pi_{1/2} J = 5/2$) inverting the ground state Λ -doublet levels responsible for the 1,667 MHz maser. These models suggest for Arp220 that the maser emission arises predominantly in very small, dense clouds, near the nucleus.

We observed the 34.97–35.54 μm spectrum of Arp220 with the Short Wavelength Spectrometer (SWS) on the Infrared Space Observatory (ISO). After processing, the spectrum (Fig. 1) has a resolving power of about 1,200. At 35.45 μm the [Si II] line (rest wavelength 34.814 μm) is seen in emission. The OH rotational transition is split into two Λ -doublet states with wavelengths of 34.629 and 34.603 μm . Hyperfine splitting of the doublet components is small (a few km s^{-1}), and is safely ignored in this analysis. The mean multiplet rest wavelength is 34.616 μm . The blended OH lines are apparent as a single resolved feature with maximum optical depth 0.21 at 35.23 μm , yielding a velocity redshift of 5,320 km s^{-1} .

To analyse the OH absorption, we derived the photon flux, and shifted the data into the rest frame, adopting a redshift^{5,10} of 5,380 km s^{-1} . We assumed that the megamaser emission and the infrared absorption are isotropic, and adopted a distance⁸ to Arp220 of 72 Mpc. We determined a baseline for the OH absorption by fitting a linear continuum, and then inverted the absorption to obtain the spectrum presented in Fig. 2, showing the photon absorption rate in the OH transition. Integrating across the line, the total absorption is 2.3×10^{55} photons s^{-1} .

The luminosity of the 1,667 MHz OH-megamaser line in Arp220 is¹⁵ 1.4×10^{36} ergs s^{-1} , assuming isotropic emission, or 1.3×10^{53} photons s^{-1} . The 1,667-MHz line accounts for $\sim 70\%$ of the total OH-megamaser emission from Arp220⁴, the rest being in lines at 1,612, 1,665 and 1,720 MHz. The radiative pump models so far constructed^{3,4} suggest that pump efficiencies can be as high as 1%, so the 35- μm absorption in Arp220 is consistent with the models, and in fact could power the observed OH masers on its own. In practice, the 53- μm transition should also play a role. This is the first time a far-infrared OH transition has been observed in a megamaser galaxy, and provides the first direct observational support for the radiative pump models.

From the strength of the OH absorption, we can estimate an OH column density of $4 \times 10^{17} \text{ cm}^{-2}$, just slightly lower than predicted by some models^{3,11}. For a single absorbing cloud of radius¹¹ 1 pc, the OH density is $\sim 0.1 \text{ cm}^{-3}$, much higher than predicted by models⁴.

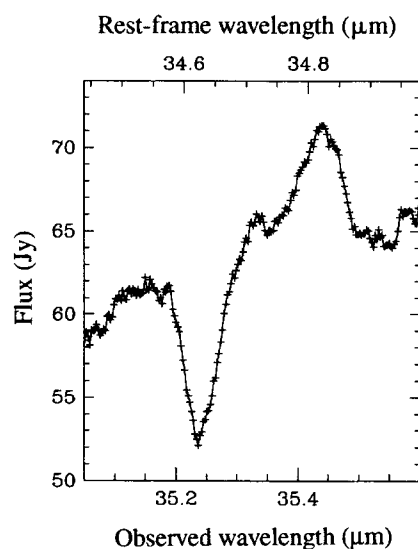


Figure 1 ISO SWS spectrum of Arp220. The emission feature at $\sim 35.45 \mu\text{m}$ is due to [Si II], and the strong absorption centred around 35.23 μm is the OH 35- μm multiplet. Both the observed wavelength (bottom scale) and the rest-frame wavelength, corrected for a 5,320 km s^{-1} redshift (top scale), are indicated.

From the OH column density we determine the OH excitation temperature⁵ to be 46 K (for an isothermal OH column), very similar to the estimate from radio maser lines⁵. For an [OH/H₂] abundance ratio of 10⁻⁶, appropriate for chemical equilibrium in warm gas^{3,12}, this yields a total gas column density of $4 \times 10^{23} \text{ cm}^{-2}$, which, assuming the mean Galactic reddening law¹³, yields a reddening $A_V \approx 200$, and a 35- μm optical depth close to 1. Other observations have been used to estimate^{14,15} $A_V \approx 50\text{--}1,000$ for Arp220. The above calculations are very crude, and subject to numerous simplifications (namely, the absorption originates from a single cloud with uniform temperature, and the dust-to-gas ratio and dust absorption properties for the absorbing cloud are similar to those for the interstellar medium in our galaxy). However, the consistency of the results with those other independent studies implies that they are at least of moderate accuracy.

From the continuum flux and the absorption depth we can make some general inferences about the absorbing regions. Models have assumed that the pumping infrared radiation comes from warm dust embedded in the maser cloud itself^{3,4}, and derived dust temperatures of $\sim 60 \text{ K}$ for a 3-pc cloud. At the line centre, 9 Jy of continuum is absorbed. Because the maximum (black-body) flux at 34.8 μm from a 60 K, 3-pc cloud at the distance of Arp220 is only 0.2 mJy, compared with the observed 9-Jy absorption, this model clearly does not work (a much larger or warmer dusty region is needed to supply the 34.8- μm photons for absorption). VLBI results¹¹ in fact reveal maser regions with sizes of $\leq 1 \text{ pc}$, with spatially coincident continuum hotspots of comparable size surrounded by a larger ($\sim 100 \text{ pc}$) diffuse source. The ISC data allow us to determine the minimum parameters of such regions.

If most of the observed 63 Jy comes from the diffuse region, at a level of 54 Jy, and the remaining 9 Jy represents the flux from a hot inner (pump) region which is then completely absorbed by OH in the diffuse region, the latter region must be at least as hot as 180 K (if it emits as efficiently as a black body), in order to supply the required 54 Jy. If the infrared-pumping cloud itself is as small as 1 pc, the 9 Jy it contributes at 35 μm , which is fully absorbed by the OH cloud, implies that its temperature is at least 21,000 K if it emits as a black body; if its size were the 3 pc suggested above, the corresponding temperature would be 2,400 K. The high colour temperature and compact nature (1 pc) of the former would imply an obscured AGN, whereas the latter temperature and size (3 pc) would be more consistent with a starburst. The complete lack of any high-ionization spectral lines, which are AGN diagnostics, in

the spectra of Arp220^{14,15} render an AGN model unlikely, unless the extinction is very much higher than we estimate above. We therefore conclude that a more consistent picture emerges from an extended starburst model. The maser emission may be beamed (that is, the coherence is directional) thereby making the region appear small in OH, and so the OH cloud may be as large as the entire starburst or very much smaller. An AGN is not required, nor is one indicated by these data, though we cannot rule out the presence of such an object.

If the OH radicals absorbing at 35 μm are pumping the radio maser lines, then the absorption should, in principle, have the same Doppler-broadened profile as the maser lines. The 35- μm OH multiplet comprises six hyperfine split lines. The separation of the dominant multiplet members is¹⁶ 0.026 μm or 225 km s^{-1} . The full-width at half-maximum (FWHM) of the line core is 490 km s^{-1} (0.058 μm) and that of the underlying pedestal is 940 km s^{-1} (0.11 μm). The 1,667-MHz maser line in Arp220 has a core with a FWHM of 130 km s^{-1} and a pedestal with a FWHM of $\sim 250 \text{ km s}^{-1}$ (refs 5, 11). Convolution of the SWS spectral response with the intrinsic line width as indicated by the maser lines indicate that we should expect a core FWHM of $\sim 430 \text{ km s}^{-1}$ and a pedestal FWHM of $\sim 680 \text{ km s}^{-1}$. Thus the pedestal width is $\sim 30\%$ larger than might be expected from the radio observations, but the core, which contains most ($\sim 80\%$) of the line flux, is as expected. The pedestal is roughly symmetric, which is consistent with the radio maser line profile except for the discrepant width. The maser gains should be greater than 1 in Arp220⁴. In a given line component (either the pedestal or the core), the maser amplification will be preferentially strengthened at the strongest velocities, 'narrowing' the OH maser line. This might explain the comparison between our ISO absorption pedestal and the maser line pedestal. Alternatively perhaps not all of the absorbing cloud generates masers, or does so along our line of sight, and the extra velocity width in absorption corresponds to unseen masers or unamplified components. The loss of half the absorbed photons to unseen masers would not affect our conclusions. Alternatively, keplerian velocities in the infrared source may be sufficient to explain the width of the absorption pedestal¹¹.

Our current ISO programme will observe OH transitions at 35, 53, 79, 96, 98, 119 and 163 μm in several megamaser galaxies and OH/IR stars (evolved red giant stars whose dense stellar winds sustain OH maser emission). These observations should provide a more comprehensive test of the accuracy of radiative pump models, by charting the course of the radiative cascade through the OH energy levels, and by determining the power absorbed in all the pumping lines. These results are unlikely to change qualitatively the results presented here, but may refine them numerically by as much as several tens of per cent. The primary remaining difficulty about the origin of the far-infrared luminosity of Arp220 is the determination of the extinction to the power source. Other studies have shown that if the extinction is as high as we have determined, then the strengths of low-excitation spectral lines are consistent with a starburst¹⁵, whereas the high-excitation lines required by an AGN are so far undetected. An unambiguous determination of the extinction could thus confirm the presence of a starburst. If, however, the extinction were shown to be much lower than we have determined, this would argue against the presence of either a starburst or an AGN, leaving the origin of the far-infrared luminosity of Arp220 highly mysterious. □

Received 20 September 1996; accepted 4 March 1997.

1. Sanders, D. B. *et al.* Ultraluminous infrared galaxies and the origin of quasars. *Astrophys. J.* **325**, 74–91 (1988).
2. Solomon, P. M. & Sage, L. J. Star-formation rates, molecular clouds, and the origin of the far-infrared luminosity of isolated and interacting galaxies. *Astrophys. J.* **334**, 613–625 (1988).
3. Henkel, C., Güsten, R. & Baan, W. A. Rotationally excited OH in megamaser galaxies. *Astron. Astrophys.* **185**, 14–24 (1987).
4. Randell, J., Field, D., Jones, K. N., Yates, J. A. & Gray, M. D. The OH zone in OH megamaser galaxies. *Astron. Astrophys.* **300**, 659–674 (1995).
5. Baan, W. A., Wood, P. A. D. & Hashchick, A. D. Broad hydroxyl emission in IC 4533. *Astrophys. J.* **260**, L49–L52 (1982).
6. Baan, W. A. in *Astrophysical Masers* (eds Clegg, A. W. & Nedoluha, G. E.) 73–82 (Springer, Berlin, 1993).

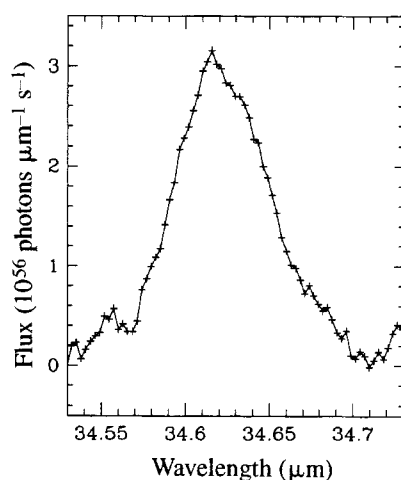


Figure 2 The strength of the OH absorption, expressed as absolute flux at the source assuming isotropic emission. The integrated absorption strength from this spectrum is $2.28 \times 10^{55} \text{ photons s}^{-1}$.

7. Baan, W. A. Infrared properties of OH galaxies. *Astrophys. J.* **338**, 804–811 (1989).
8. Graham, J. R. *et al.* The double nucleus of Arp 330 unveiled. *Astrophys. J.* **354**, L5–L8 (1990).
9. Baan, W. A. & Haschick, A. D. Nuclear antics of IC 4553 (Arp 220). *Astrophys. J.* **454**, 745–748 (1995).
10. Diamond, P. J., Norris, R. P., Baan, W. A. & Booth, R. S. VLB observations of the OH megamaser in IC 4553 (Arp 220). *Astrophys. J.* **340**, L49–L52 (1989).
11. Lonsdale, C. J., Diamond, P. J., Smith, H. E. & Lonsdale, C. J. Compact OH megamaser and probable quasar activity in the galaxy Arp 220. *Nature* **370**, 117–120 (1994).
12. Goldreich, P., Scoville, N. OH/IR Stars I—Physical properties of circumstellar envelopes. *Astrophys. J.* **205**, 144–154 (1976).
13. Savage, B. D. & Mathis, J. S. Observed properties of interstellar dust. *Annu. Rev. Astron. Astrophys.* **17**, 73–111 (1979).
14. Sturm, E. *et al.* ISO-SWS spectroscopy of Arp 220: a highly obscured starburst galaxy. *Astron. Astrophys.* **315**, L133–L136 (1996).
15. Lutz, D. *et al.* What powers luminous infrared galaxies? *Astron. Astrophys.* **315**, L137–L140 (1996).
16. Destombes, J. L., Marliere, C., Baudry, A. & Brillet, J. The exact hyperfine structure and Einstein A-coefficients of OH—consequences in simple astrophysical models. *Astron. Astrophys.* **60**, 55–60 (1977).

Acknowledgements. This work is based on observations with ISO, an ESA project with instruments funded by the ESA Member States (especially the PI countries: France, Germany, the Netherlands and the UK) and with the participation of ISAS and NASA.

Correspondence and requests for materials should be addressed to C.J.S. (e-mail: skinner@stsci.edu).

Individual single-wall carbon nanotubes as quantum wires

Sander J. Tans⁺, Michel H. Devoret[†], Hongjie Dai[‡], Andreas Thess[‡], Richard E. Smalley[‡], L. J. Geerligs^{*} & Cees Dekker^{*}

^{*} Department of Applied Physics and DIMES, Delft University of Technology, Lorentzweg 1, 2628 CJ Delft, The Netherlands

[‡] Center for Nanoscale Science and Technology, Rice Quantum Institute, and Departments of Chemistry and Physics, MS-100, Rice University, PO Box 1892, Houston, Texas 77251, USA

Carbon nanotubes have been regarded since their discovery¹ as potential molecular quantum wires. In the case of multi-wall nanotubes, where many tubes are arranged in a coaxial fashion, the electrical properties of individual tubes have been shown to vary strongly from tube to tube^{2,3}, and to be characterized by disorder and localization⁴. Single-wall nanotubes^{5,6} (SWNTs) have recently been obtained with high yields and structural uniformity⁷. Particular varieties of these highly symmetric structures have been predicted to be metallic, with electrical conduction occurring through only two electronic modes^{8–10}. Because of the structural symmetry and stiffness of SWNTs, their molecular wavefunctions may extend over the entire tube. Here we report electrical transport measurements on individual single-wall nanotubes that confirm these theoretical predictions. We find that SWNTs indeed act as genuine quantum wires. Electrical conduction seems to occur through well separated, discrete electron states that are quantum-mechanically coherent over long distance, that is at least from contact to contact (140 nm). Data in a magnetic field indicate shifting of these states due to the Zeeman effect.

The electronic band structure of carbon nanotubes presents some peculiar features. Owing to the hexagonal symmetry of a graphene sheet, the Fermi contour expected for a two-dimensional lattice reduces to six points¹¹. In a nanotube, the component of momentum along the circumference of the tube will be quantized owing to periodic boundary conditions. For so-called ‘armchair’ SWNTs (see ref. 12 for an overview), this results in only two one-dimensional gapless modes of propagation parallel to the tube axis^{8–10}. In the simplest approach, we consider the electron states to be four-fold degenerate, including spin. The strictly one-dimensional nature of the modes implies that the angular momentum contribution to the magnetic energy is zero. The magnetic energy is expected to be

dominated by the spin contribution $g\mu_B B$ (Zeeman energy), where μ_B is the Bohr magneton, g is the Landé g -factor, and B the magnetic field. For a finite one-dimensional tube of length L , the component of momentum along the axis will also be quantized. This results in ‘particle-in-a-box’ energy separation between levels of $\Delta E = h v_F / 2L$. From band-structure calculations¹¹, we can find an estimate for the Fermi velocity $v_F \approx 8.1 \times 10^5 \text{ m s}^{-1}$. Even a 3- μm -long tube would thus have a considerable splitting energy of $\Delta E \approx 0.6 \text{ meV}$.

The nanotubes are produced by laser-vaporization of carbon with an admixture of Ni/Co, as described elsewhere⁷. This technique results in samples enriched in armchair SWNTs with a uniform diameter of 1.38 nm. After deposition on the Si/SiO₂ substrate, the tubes are imaged by an atomic force microscope (AFM). The images show tubular features with lengths of the order of micrometres and heights between ~ 1 and $\sim 50 \text{ nm}$. The 1-nm-high features are identified as individual single-wall tubes. Discrepancy of this value with the assumed diameter of 1.38 nm is believed to come from plastic deformation of the tube resulting from tip-sample forces. The larger features in the images are ropes of parallel tubes⁷. Figure 1 shows an AFM image of an individual tube connected to two electrodes and a corresponding circuit diagram. The two-point resistance at room-temperature of single-tube samples generally is $\sim 1 \text{ M}\Omega$, and appears to scale with the overlap area with the electrode. (The resistance at room temperature is 550 k Ω for the sample in Fig. 1.) On a different single-tube sample with a four-point geometry, we determined the contact resistance to be around 300 k Ω at room temperature and 1 M Ω at 4 K. The measurements presented here were performed on the sample displayed in Fig. 1 in a dilution refrigerator with a mixing-chamber base temperature of 5 mK. Three other samples yielded similar results.

Typical current–voltage curves at various gate voltages are shown in Fig. 2a. Around zero bias voltage a clear gap is present. For higher voltages, the current increases in steps. The gap appears to be suppressed for certain gate voltages and has a maximum value in between (see inset). In Fig. 2b, the current is shown as a function of the gate voltage at bias voltage $V_{\text{bias}} = 30 \mu\text{V}$. The sequence of current peaks indicates that the suppression of the gap occurs in a

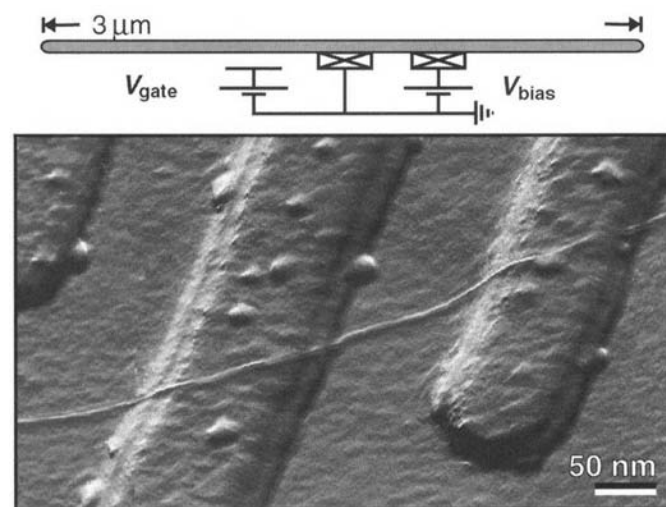


Figure 1 AFM tapping-mode image of a carbon nanotube on top of a Si/SiO₂ substrate with two 15-nm-thick Pt electrodes, and a corresponding circuit diagram. The nanotubes are deposited by spin-coating of a drop of nanotube suspension. This tube has a diameter of $\sim 1 \text{ nm}$, as deduced from AFM height profiles, and is identified as an individual single-wall nanotube. The total length of the tube is 3 μm , with a section of 140 nm between the contacts to which a bias voltage V_{bias} is applied. A gate voltage V_{gate} applied to the third electrode in the upper-left corner of the image is used to vary the electrostatic potential of the tube.

[†] Permanent address: Service de Physique de l'Etat Condensé, CEA-Saclay, F-91191 Gif-sur-Yvette, France.

quasi-periodic fashion. The same pattern of peaks persists up to very large scans of gate voltage V_{gate} (at least from -2.5 to $+2.5$ V). The observations indicate Coulomb charging¹³ of the nanotube. Coulomb charging occurs when the contact resistance is larger than the quantum resistance R_K ($R_K = h/e^2 \approx 26$ k Ω) and when the total capacitance C of the object is so small that adding a single electron costs a significant charging energy $E_c = e^2/2C$. At low temperatures, that is, $k_B T \ll E_c$, the current is thus blocked and will flow only when $V_{\text{bias}} > e/2C$. However, zero-bias conductance can be restored by precisely 'tuning' the electrostatic potential of the tube with the gate voltage. The two traces in Fig. 2b were taken under identical conditions and show an occasional doubling of certain peaks. This bistability may be the result of switching offset charges¹³ that shift the potential of the tube. We have concentrated measurements on peaks that do not show this bistability.

In Fig. 3 we show measurements of a single conductance peak at $V_{\text{bias}} = 10$ μ V. The height of the conductance peak in Fig. 3 appears to decrease strongly with increasing temperature. This is a signature of resonant tunnelling through a discrete electron level that is aligned with the Fermi energy E_F of the electrodes (see left inset)¹⁴. We label this level as gg (ground state to ground state) because when it alternates in time between being unoccupied and occupied, the tube is in the ground state in both cases. This gg-orbital is well separated by an energy ΔE from the occupied orbital below, and by $\Delta E'$ from the unoccupied orbital above. The conductance through the gg-orbital is now only dependent on the difference in electron occupancy at E_F between the two contacts which, at resonance, goes down with increasing temperature. If the density of states in the tube were continuous, the conductance maximum would be constant with increasing temperature¹⁴. The resonant tunnelling through discrete electron levels implies that single molecular orbitals carry the current, and accordingly are phase coherent and extended over a distance of at least 140 nm (the distance between the electrodes). This is a crucial observation. It shows that the electrons are not strongly localized despite the

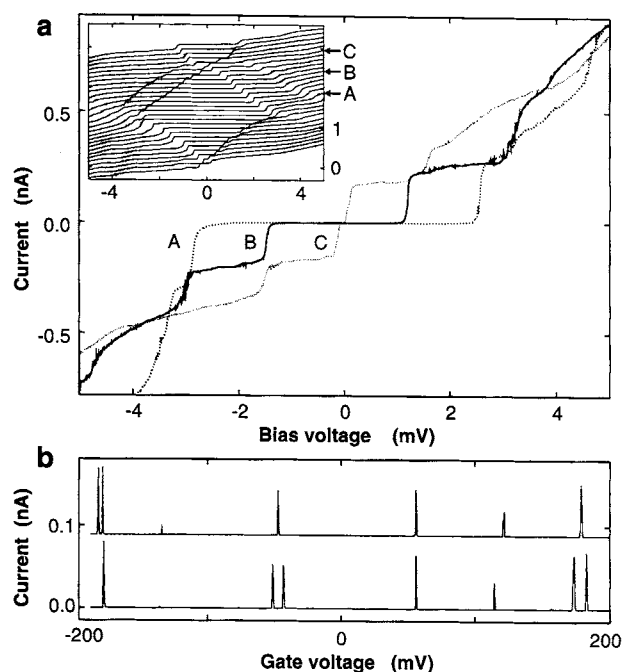


Figure 2 **a**, Current-voltage curves of the nanotube at a gate voltage of 88.2 mV (trace A), 104.1 mV (trace B) and 120.0 mV (trace C). Inset, more $-V_{\text{bias}}$ curves with V_{gate} ranging from 50 mV (bottom curve) to 136 mV (top curve), with vertical offsets for clarity. The variation with V_{gate} of the gap around zero bias voltage implies Coulomb charging of the tube. The stepwise increase of the current at higher voltages may result from an increasing number of excited states entering in the bias window. **b**, Current versus gate voltage at $V_{\text{bias}} = 30$ μ V. Two traces are shown that were performed under the same conditions.

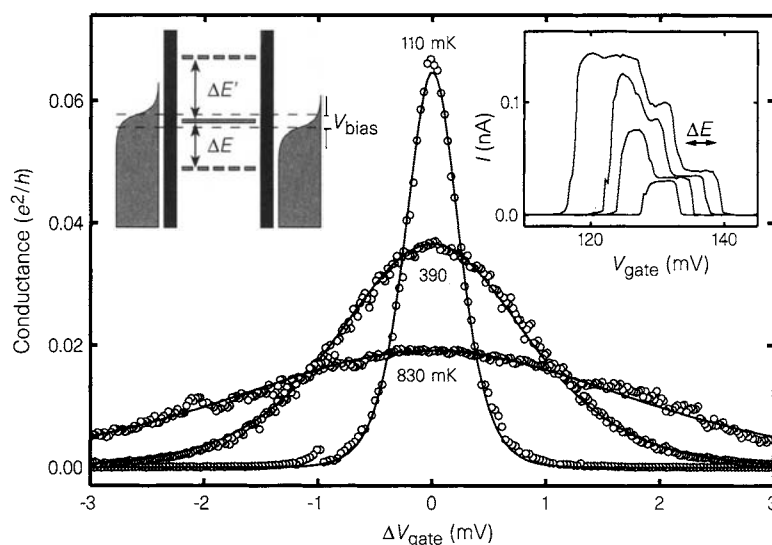


Figure 3 Conductance $G = I/V_{\text{bias}}$ versus ΔV_{gate} at low bias voltage $V_{\text{bias}} = 10$ μ V and different temperatures. Solid lines are fits of $G \propto \cosh^{-2}(e\Delta V_{\text{gate}}/\alpha 2k_B T)$, corresponding to the model of a single molecular level that is weakly coupled to two electrodes¹⁴ (see inset). The factor α converts ΔV_{gate} into the corresponding electrostatic potential shift of the tube. For this peak $\alpha = 16$. The solid bar in the left inset denotes the gg-level (see text) and is well separated by energies ΔE , $\Delta E' \gg k_B T$ from the eg (lower broken bar) and ge (upper broken bar) levels that involve excited states of the tube. At a certain gate voltage, the gg-level will be in the bias window, resulting in resonant tunnelling of electrons. ΔV_{gate} is the gate

voltage taken relative to V_{gate} at resonance. The indicated temperatures result from fits of this model. The corresponding thermometer temperatures are 5, 240 and 600 mK, respectively. The discrepancy between the two temperatures can be attributed to additional broadening due to the finite bias voltage (10 μ V ≈ 116 mK) and residual noise in the measurement system. From the maximum of the peak we obtain a lower bound on electronic transmission of $1/60$, which includes the contact resistance. In the right inset we plot the current as a function of V_{gate} for higher $V_{\text{bias}} = 0.4, 0.8, 1.2$ and 1.6 mV. For this peak, $\alpha = 12$. The steps in the peak are a result of lower-lying levels in the bias window.

one-dimensional nature of the tube and potentially perturbing influences like structural imperfections, bends and twists of the tube, and charge defects in the SiO₂ support. Apparently the structural symmetry and stiffness of the molecule does indeed result in robustness of the phase coherence of the molecular orbitals.

When the conductance peaks are measured at higher bias voltage, step-like structures are observed (Fig. 3, right inset). These structures are attributed to contributions from additional levels. Upon reducing the gate voltage, a first current plateau develops when the gg-level enters the bias-window from below. When the levels are pushed up further in energy by ΔE , electrons can now also leave the tube from the lower-lying electron level, leading to a higher current plateau. This lower-lying level is labelled as eg because when it is emptied, the tube is left in an excited state. Current will be blocked again when the gg-level finally leaves the bias-window. From the total peak width, which is set by V_{bias} , we deduce that 12-mV change in V_{gate} results in a change of tube potential of about 1 mV. This conversion can be understood in terms of voltage division: the gate capacitance is one-twelfth of the total capacitance of the tube. The width of the rightmost plateau corresponds to ΔE . On average, we find $\Delta E \approx 0.4$ meV. This value is consistent with the order-of-magnitude estimate $\Delta E \approx 0.6$ meV for a 3- μm -long tube, as mentioned above. This in fact suggests that the coherence extends over the full tube length of 3 μm .

Returning to Fig. 2: the stepwise current increase with increasing V_{bias} may also result from an increasing number of levels within the bias window. These steps are not observed for samples with ropes of tubes. Because of their larger size, they are expected to have a much lower energy separation between levels. A Coulomb staircase¹⁵, which occurs when there is a large difference between the two contact resistances, does not seem to be the origin of the observed steps. For example, the 2-mV width of the step of curve B between $V_{\text{bias}} = 1.2$ and 3.2 mV should then be equal to the width of the

maximum gap (~ 6 mV). From Fig. 2 we can also determine the Coulomb charging energy E_c as the width of the maximum gap corresponds to $2(E_c + \Delta E)$. The resulting $E_c \approx 2.6$ meV is in good agreement with an estimate of $E_c \approx 2.5$ meV based on the geometrical self-capacitance of a 3- μm -long tube (3×10^{-17} F). Summing up, the tube appears to have a discrete energy spectrum with an energy level separation of ~ 0.4 meV, and an extra gap is observed due to a Coulomb energy of ~ 2.6 meV. Electron transport through discrete electron states has been found previously in systems based on semiconductor two-dimensional electron gases¹⁵ and recently also in a metal system¹⁶.

Single-electron levels will shift in energy upon application of a magnetic field. We measure ΔE as a function of magnetic field B by following the bottom trace of the right inset of Fig. 3 (see Fig. 4). The gg and eg levels appear to move toward each other (situation B in Fig. 4), cross, and separate again (situation C). From the slope of the diagonal line in the contour-plot $d(\Delta E_c)/d(\mu_B B)$, we find a g -factor of 1.9 ± 0.2 . This is consistent with magnetic-resonance data⁷ and suggests no spin-orbit coupling. It appears that the Zeeman energy is dominant, confirming the expected negligible angular-momentum term. Note that the data do not show splitting of states at $B = 0$ T, as would be expected in the simplest model of level filling and assuming spin degeneracy. This deviation is not yet understood. The plateaux in Fig. 4 appear to be flat over a large range of magnetic field. The tubes thus do not exhibit a magnetoresistance, which would be manifested by a gradual decrease of the peak amplitude as a function of magnetic field. This provides an additional indication that the electronic states are indeed extended and have a one-dimensional nature.

The measurements reported here are, to our knowledge, the first measurements of the electrical transport properties of individual SWNTs. Remarkably, these molecules appear to behave as coherent quantum wires. Considering the strongly one-dimensional nature

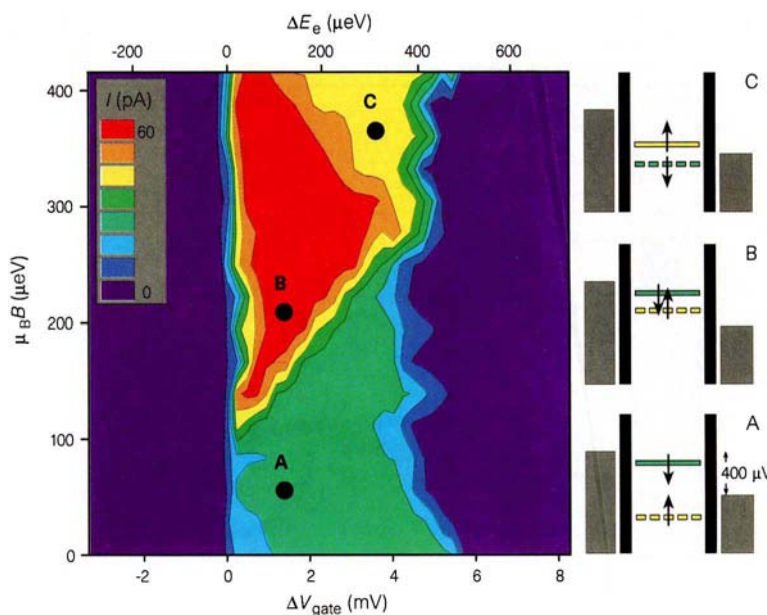


Figure 4 Main figure, current level (see inset colour scale), as a function of $\mu_B B$ and ΔV_{gate} and corresponding relative electrostatic energy ΔE_e , at $V_{\text{bias}} = 400$ μV . The magnetic field is applied perpendicular to the tube axis and has a maximum value of 7.5 T. The left current edge has been aligned for all B -values. At low magnetic fields there is a single 'green' current plateau, resulting from the 'green' gg-level with, say, spin down (situation A, right-hand diagrams). The nearest level is the eg-level (yellow) which has spin up and is already outside the bias window. Near 3.5 T (situation B) the two levels appear to have shifted towards each other.

They now both contribute to a higher, 'red' current level. At $\Delta V_{\text{gate}} \approx 2.5$ mV, the eg-level leaves the bias window. The width of the remaining green plateau corresponds to ΔE . Continuing this trend, the two levels cross and reverse their order (situation C). Now the 'yellow' level has become the gg-level and will leave the bias window last. These observations are consistent with shifting of states due to the Zeeman effect. Note that the difference between the 'green' and 'yellow' current levels implies a different coupling of the two corresponding orbitals to the electrodes.

of the conducting modes in our SWNTs, deviations from Fermi-liquid behaviour may be expected; further analysis and measurements may address this issue. □

Received 3 December 1996; accepted 11 March 1997.

1. Iijima, S. Helical microtubules of graphitic carbon. *Nature* **354**, 56–58 (1991).
2. Dai, H., Wong, E. W. & Lieber, C. M. Probing electrical transport in nanomaterials: conductivity of individual carbon nanotubes. *Science* **272**, 523–526 (1996).
3. Ebbesen, T. W. *et al.* Electrical conductivity of individual carbon nanotubes. *Nature* **382**, 54–56 (1996).
4. Langer, L. *et al.* Quantum transport in a multiwalled carbon nanotube. *Phys. Rev. Lett.* **76**, 479–482 (1996).
5. Bethune, D. S. *et al.* Cobalt-catalysed growth of carbon nanotubes with single-atomic-layer walls. *Nature* **363**, 605–607 (1993).
6. Iijima, S. & Ishihashi, T. Single-shell carbon nanotubes of 1-nm diameter. *Nature* **363**, 603–605 (1993).
7. Thess, A. *et al.* Crystalline ropes of metallic carbon nanotubes. *Science* **273**, 483–487 (1996).
8. Mintmire, J. W., Dunlap, B. I. & White, C. T. Are fullerene tubules metallic? *Phys. Rev. Lett.* **68**, 631–634 (1992).
9. Hamada, N., Sawada, A. & Oshiyama, A. New one-dimensional conductors: graphitic microtubules. *Phys. Rev. Lett.* **68**, 1579–1581 (1992).
10. Saito, R., Fujita, M., Dresselhaus, G. & Dresselhaus, M. S. Electronic structure of chiral graphene tubules. *Appl. Phys. Lett.* **60**, 2204–2206 (1992).
11. Wallace, P. R. The band theory of graphite. *Phys. Rev.* **71**, 622–633 (1947).
12. Dresselhaus, M. S., Dresselhaus, G. & Eklund, P. C. *Science of Fullerenes and Carbon Nanotubes* (Academic, San Diego, 1996).
13. Grabert, H. & Devoret, M. H. (eds) *Single Charge Tunneling* (Plenum, New York, 1992).
14. van Houten, H., Beenhakker, C. W. & Staring, A. A. M. in *Single Charge Tunneling* (eds Grabert, H. & Devoret, M. H.) 167–216 (Plenum, New York, 1992).
15. Geerligs, L. J., Harmans, C. J., P.M. & Kouwenhoven, L. P. (eds) *The Physics of Few-Electron Nanostructures* (North-Holland, Amsterdam, 1993).
16. Ralph, D. C., Black, C. T. & Tinkham, M. Spectroscopic measurements of discrete electronic states in single metal particles. *Phys. Rev. Lett.* **74**, 3241–3244 (1995).

Acknowledgements. We thank I. Kouwenhoven, T. Oosterkamp, R. Groeneveld and Y. Nazarov for discussions and B. van den Enden for technical assistance. The work was supported by the Dutch Foundation for Fundamental Research on Matter (FOM).

Correspondence and requests for materials should be addressed to C.D.

Control of spatial orientation and lifetime of scroll rings in excitable media

Michael Vinson*†, Sergey Mironov*‡, Scott Mulvey* & Arkady Pertsov*

* Department of Pharmacology, SUNY Health Sciences Center, Syracuse, New York 13210, USA

† Department of Physics, Shippensburg University, Shippensburg, Pennsylvania 17257, USA

‡ Institute of Cellular Biophysics, Puschino, Moscow Region 142292, Russia

Excitable media, which range from autocatalytic chemical systems^{1–3} to biological cells and tissues^{4–7}, can maintain organized structures in the form of rotating spiral waves of excitation^{8–11}. The dynamics of spiral waves in two-dimensional systems have been shown to be susceptible to control by external fields (such as electric, thermal and optical)^{12–14}. In three dimensions, the analogues of spiral waves are scroll waves^{9,15,16}. Here we show that an external field—a temperature gradient—can be used to control a particular class of scroll waves called scroll rings. The gradient allows scroll rings to be precisely oriented in space, and their spontaneous shrinkage to be accelerated, decelerated or even reversed (so that the ring expands). The temperature gradient also influences the lifetimes of the scroll rings. We suggest that these dynamics are likely to be generic to other types of field gradients and other excitable media.

The control of three-dimensional (3D) scroll waves is a more complex problem than the control of their two-dimensional analogues, spiral waves. The dynamics of a 3D scroll wave are determined not only by the properties of the excitable medium but also by the shape of the filament (the vortex line about which the wave rotates). Depending on the topology of the filament (for example, open,

closed loop, knot), as well as on its curvature, torsion and twist, the dynamics of scroll waves can be significantly different even in the same excitable medium^{11,17,18}. This implies that the possibility of controlling scroll waves should strongly depend on a given filament configuration.

Our experiments were carried out in a gel-immobilized Belousov–Zhabotinsky (BZ) reaction^{19,20}. Figure 1a is a diagram of a cross-section of a scroll ring, showing the circular axis (filament) about which the wave rotates. Physically, the filament is characterized by zero amplitude of the excitation wave²¹. The wave has the shape of a toroidal scroll, which in this cross-section appears as two counter-rotating spirals. The spatial orientation of the ring is described by the unit vector **S**, normal to the ring plane and with orientation uniquely determined by the sense of rotation of the ring. Figure 1b shows a photograph of an experimental scroll ring in the BZ reaction; Fig. 1c is a reconstruction of its filament, appearing in the projection as a dark ellipse.

Our observations show that there are two effects of an applied temperature gradient: it reorients the plane of the scroll ring so as to align the vector **S** with the gradient vector **g**, and it significantly affects the lifetime of the ring. Once the vectors **S** and **g** become parallel, they remain parallel for the duration of the experiment. (Experiments with initial angles of 0° and 180° showed no reorientation.) The observed relaxation time of the reorientation was found to be long compared to the characteristic rotation period, t_s , of the spiral around the filament.

Figure 2a shows the reorientation of a scroll ring in a gradient of 10 °C cm^{−1}. The filament is shown in lateral projection, and appears as a dark ‘dumb-bell’ that slowly rotates in the anticlockwise direction. The angles θ between **S** and **g** are shown as filled circles in Fig. 2b. The effect of **g** on the lifetime of the scroll ring can be evaluated by comparing the dynamics of the filament radius (r) in the presence and absence of gradient (Fig. 2b, bottom). In the absence of gradient, the ring shrinks and collapses in 20 minutes (open circles). The gradient prolongs the lifetime of the ring to almost 70 minutes which is about 3.5 times greater than the lifetime in the absence of the field. In a stronger gradient, we found that the

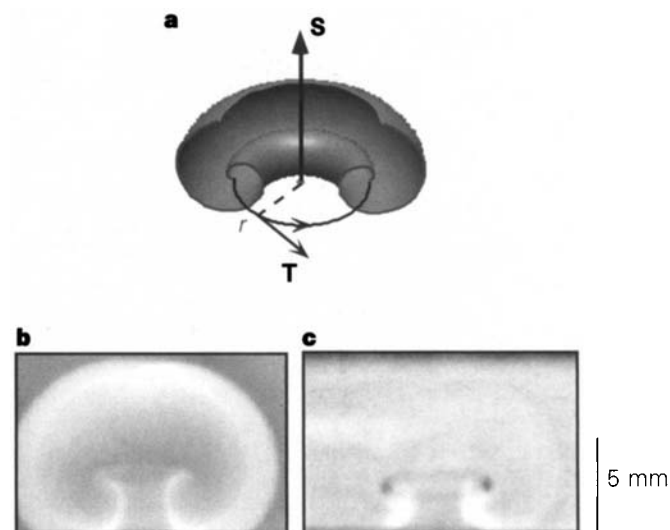


Figure 1 a, Schematic cross-section of a scroll ring. The dark circle of radius r in the centre is the filament of the ring, about which the wave rotates (the spiral on the right rotates clockwise). The unit tangent vector **T** is in the direction of the local angular velocity. The vector **S** is normal to the ring plane, with orientation given by applying the right-hand rule to **T** along the ring. b, Photograph of lateral projection of a tilted scroll ring in a gel-immobilized BZ reaction. Excited (oxidized) regions appear lighter. The angle between vector **S** and the vertical axis is 11°. c, Photograph of the filament in the same projection. The filament is roughly circular, appearing elliptical in the figure due to the projection.

shrinking may be reversed and the ring may actually expand. Figure 2c shows the expansion effect in a gradient of $20^{\circ}\text{C cm}^{-1}$. As soon as the gradient was removed (arrow) the expansion stopped and ring rapidly collapsed. The expansion effect was maximal when $\theta = 0$.

By contrast, when $\theta = 180^{\circ}$ the gradient accelerated the shrinking. For rings of the same initial size and gradient as in the experiment of Fig. 2a, b, the rings collapsed in less than 8 minutes. The last value is comparable with the rotation cycle of the spiral, t_s . At temperatures of $\sim 20^{\circ}\text{C}$ (the temperature in the centre of cuvette in our experiments) t_s is ~ 7 min.

The effects of an external field are reproducible in computer simulations of the 3D BZ reaction, in a broad parameter range. Figure 3 shows reorientation and shrinking of the ring filament in the simulations, with a small temperature gradient ($4^{\circ}\text{C cm}^{-1}$) and in a domain of twice the size of the experimental system. Expansion of the ring (at larger gradients) and acceleration of shrinking (when $\theta = 180^{\circ}$) were also reproduced in simulations.

The physical mechanism of the effect of temperature gradient on a scroll ring may be understood by combining the curvature-induced motion of the filament, N (which determines the ring dynamics in the absence of a gradient^{11,18,22}) with the gradient-induced drift velocity G of the filament¹⁹. The vector N always points towards the centre of the ring and its magnitude is proportional to the curvature of the ring, $1/r$. In the BZ reaction, G is observed to be perpendicular to both the gradient and the filament^{19,23}: $G = \beta T \times g$, where T is the local unit tangent vector to the filament, and β is a parameter of the medium. Figure 4 shows

the result of combining these two velocities for rings with different initial orientations. The vector G always points in opposite directions at opposite points on the scroll ring, causing a net 'torque' (when $\theta \neq 0, 180^{\circ}$) that reorients the plane of the ring (Fig. 4a, b). Regardless of the initial orientation, symmetry implies that the effect of the gradient is always to reduce θ towards zero.

When θ equals zero (Fig. 4c) G opposes N and, as in the experiment, a sufficiently strong gradient causes the ring to expand. There is thus a critical gradient, depending on the curvature ($1/r$), at which $G = N$ and the ring is in equilibrium. This equilibrium state is unstable; if the gradient is slightly less than critical, then the ring shrinks and its curvature increases, causing an increase in N and further shrinking. When $\theta = 180^{\circ}$ (Fig. 4d) the gradient velocity adds to the normal velocity, accelerating the shrinking. This orientation is directionally unstable, as can be seen in Fig. 4b.

The schematic approach discussed above leads to a simple quantitative model. By adding the equation for G to the lowest-order terms of the theory of local filament dynamics^{11,18,22,24}, and specializing to the case of a circular, planar scroll ring of radius r and inclination angle θ , the following equations may be derived:

$$\begin{aligned} dr/dt &= -\alpha/r - \beta g \cos \theta \\ d\theta/dt &= (\beta g/r) \sin \theta \end{aligned} \quad (1)$$

where α is the proportionality constant between N and the filament curvature, estimated from the observed rate of shrinking in the absence of a gradient to be $0.12 \pm 0.01 \text{ mm}^2 \text{ min}^{-1}$. The parameter

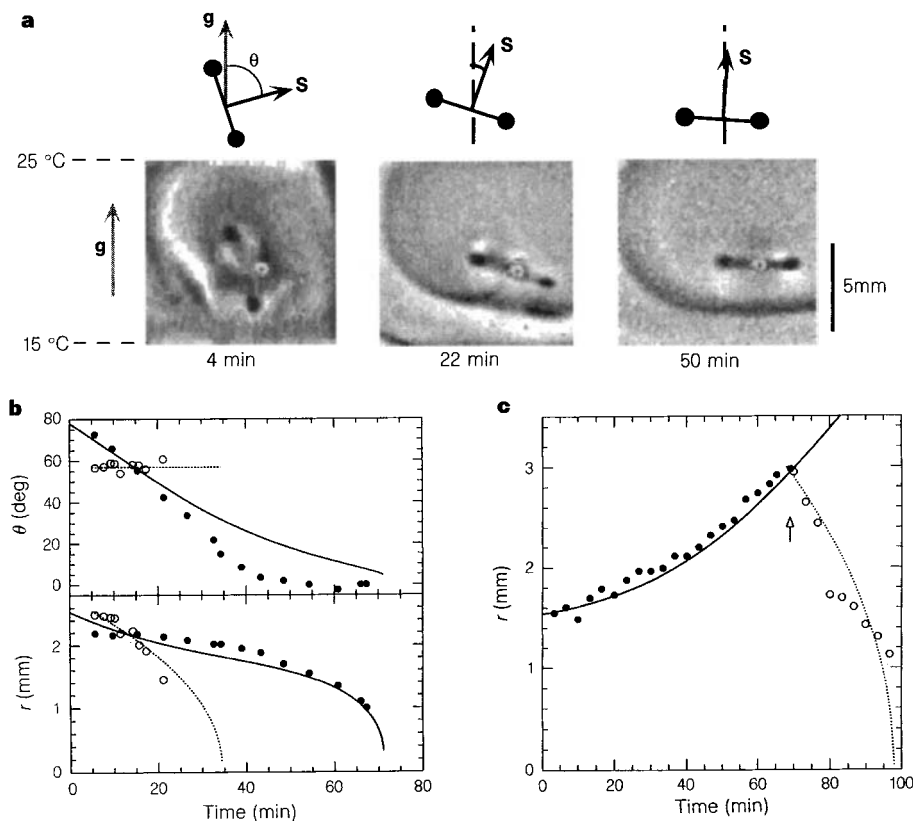


Figure 2 a, Reorientation of a scroll ring filament in a gel-immobilized BZ reaction, with external field provided by a temperature gradient of $10^{\circ}\text{C cm}^{-1}$. The gradient vector g points upwards; θ is the angle between g and S . Insets on the top show the filament locations as abstracted from the photographs. The filament appears as a dark 'dumb-bell' shape as a result of the greater optical density at the edges of the circle in lateral projection. The defect in the middle of each 'dumb-bell' is a bubble in the foreground. **b**, Plot of θ and r , the ring radius, as functions of time, from the same experiment (filled circles) and a control experiment with a constant

temperature (open circles). The solid lines are theoretical curves computed from equations (1) (see text). The dotted lines are theoretical curves for zero gradient. **c**, Expansion of a scroll ring in an experiment with large gradient ($20^{\circ}\text{C cm}^{-1}$), and initial $\theta = 0$. The arrow indicates the time when the gradient was removed. The filled circles are experimental points before removal of the gradient; open circles are after. As in **b**, the solid and dotted lines are theoretical curves computed from equations (1). The deviation of experimental points (open circles) from the theoretical curve is a result of gradient relaxation that the model does not account for.

β was found experimentally to be 0.048 ± 0.005 mm per $^{\circ}\text{C}$ per min (ref. 19). The binormal component of the curvature-induced velocity is neglected, as it is found to be very small in the BZ reaction²⁵.

Equations (1) yield qualitatively correct predictions, with the right order of magnitude of the quantities involved, as can be seen from the curves in Fig. 2b and c. However, the predicted reorientation time is greater than observed in the experiment (Fig. 2b), and the gradient used in the theoretical curves are different from the experimental values by as much as $5^{\circ}\text{C cm}^{-1}$. The discrepancies between the model and the experiment are probably due to inherent limitations of the model, specifically the assumption of small filament curvature²² and the failure to include nonlocal effects¹⁹. The latter limitation is particularly acute when θ is large, so that the part of the ring in the hotter region of the medium rotates at a faster rate, producing wavefronts that interact with the colder (more slowly rotating) part of the ring, inducing an additional filament drift that is not accounted for in equations (1).

The behaviour of a scroll ring in an external gradient is reminiscent of a current loop in a magnetic field. In the latter case, the magnetic field produces a torque on the loop, in a direction so as to align the magnetic-moment vector (which is perpendicular to the plane of the loop) with the magnetic field. There is also a force per unit length on the current loop directed outward (when $0 < \theta < 90^{\circ}$), or inward (when $90^{\circ} < \theta < 180^{\circ}$), analogous to the effect of a temperature gradient on the shrinking of a scroll ring.

Our study indicates that control over the life span of a scroll ring can be achieved by choosing proper strength and orientation of the field. By contrast, there is no such possibility of controlling the life span of two-dimensional spirals. The analysis presented above

shows that the mechanism of control is universal, and, with some minor modifications, should hold for any excitable medium and gradient. External fields, such as electric field¹³ or the temperature gradient described above, always break the isotropy of the medium resulting in a chirality-dependent filament drift. They should therefore have a similar effect on the scroll ring motion. In some systems (in which the gradient-induced drift velocity is not perpendicular to the gradient^{13,26,27}), one may anticipate an additional translational motion of the ring along its axis of symmetry. The phenomena of reorientation and expansion will still occur. $\square$

Methods

The BZ recipe was: 0.05 M NaBrO₃, 0.2 M H₂SO₄, 0.05 M CH₂(COOH)₂ and 6.25×10^{-4} M ferroin, immobilized in a 1% agarose gel. The ferroin caused excited regions to appear blue, allowing video recording using a Cohu-6500 CCD video camera equipped with a band-pass filter centred at 490 nm. The experiments were performed in $10 \times 10 \times 50$ mm cuvettes; the temperature gradient was applied by placing the cuvettes between two steel heat exchangers connected to two thermostat-controlled circulating baths. Initiation was performed by perforating the wavefront with a polyethylene tube or a beam of light of radius 2.5–3 mm. The filament was obtained by the method of video averaging, in which sequential video frames are digitally averaged, as described in ref. 19. The shape of the filament in three dimensions was determined by analysing two orthogonal projections of the vortex pattern.

Received 29 July 1996; accepted 12 February 1997.

1. Zhabotinsky, A. M. & Zaikin, A. N. *Oscillatory Processes in Biological and Chemical Systems* Vol. 2 (Nauka, Pushchino, Russia, 1971).
2. Winfree, A. T. Spiral waves of chemical activity. *Science* **175**, 634–636 (1972).
3. Müller, S. C., Plesser, T. & Hess, B. *Science* **230**, 661–663 (1985).
4. Davidenko, J. M., Pertsov, A. M., Salomonsz, R., Baxter, W. & Jalife, J. Stationary and drifting spiral waves of excitation in isolated cardiac muscle. *Nature* **335**, 349–351 (1992).

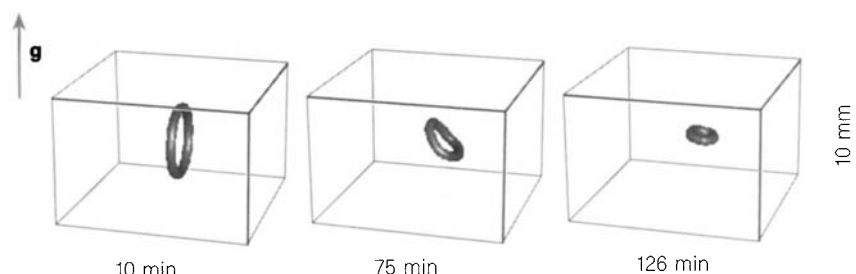


Figure 3 Scroll ring filament from a computer simulation of the BZ reaction. The gradient was $4^{\circ}\text{C cm}^{-1}$ and initial $\theta = 90^{\circ}$. Both the reorientation of the ring and its shrinking can be seen. The simulations were done using the BZ model of Aliev and Rovinsky²³, with proton concentration (h_0) of 0.15, and diffusion coefficient 2×10^{-5} cm² s⁻¹. The gradient was introduced by scaling the reaction constants using the Arrhenius (exponential) law with a 20° decrease in temperature causing

a 10-fold decrease in each rate constant. The numerical parameters were: the time step $h_t = 0.1$ s, the space step $h_x = 0.03125$ cm, and the lattice was $64 \times 64 \times 64$, corresponding to a medium size of $2 \times 2 \times 2$ cm. The computations were implemented in Fortran90 on the DECmpp massively parallel computer at the Northeast Parallel Architectures Center at Syracuse University.

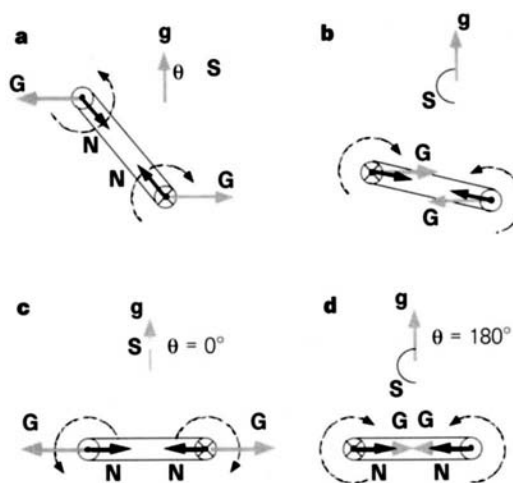


Figure 4 Schematic showing the effect of the gradient-induced velocity for different orientations of the ring. The dashed arcs indicate the direction of the scroll rotation. The vector **G** is the gradient-induced velocity, **S** is the unit normal to the ring plane, **N** is the curvature-induced normal velocity, and **g** is the gradient vector. **a**, Ring at small angle. **b**, Angle close to 180° . **c**, Zero angle. **d**, Angle of 180° .

5. Siegert, F. & Weijer, C. Analysis of optical density wave propagation and cell movement in the cellular slime mold dictyostelium discoideum. *Physica D* **49**, 224–232 (1991).
6. Gorelova, N. A. & Bures, J. Spiral waves of spreading depression in the isolated chicken retina. *Neurobiology* **14**, 353–363 (1983).
7. Lechleiter, J., Girard, S., Peralta, E. & Clapham, D. Spiral calcium wave propagation and annihilation in *Xenopus laevis* oocytes. *Science* **252**, 123–126 (1991).
8. Krinsky, V. I. (ed.) *Self-organization: Autowaves and Structures far from Equilibrium* (Springer, Berlin, 1984).
9. Winfree, A. T. & Strogatz, S. H. Organizing centers for three-dimensional chemical waves. *Nature* **311**, 611–615 (1984).
10. Swinney, H. L. & Krinsky, V. I. (eds.) *Waves and patterns in chemical and biological media*. *Physica D* **49** (1991).
11. Keener, J. P. & Tyson, J. J. The dynamics of scroll waves in excitable media. *SIAM Rev.* **34**, 1–39 (1992).
12. Schütze, J., Steinbock, O. & Müller, S. C. Forced vortex interaction and annihilation in an active medium. *Nature* **356**, 45–47 (1992).
13. Steinbock, O., Schütze, J. & Müller, S. C. Electrical-field-induced drift and deformation of spiral waves in an excitable medium. *Phys. Rev. Lett.* **68**, 248–251 (1992).
14. Steinbock, O., Zykov, V. & Müller, S. C. Control of spiral-wave dynamics in active media by periodic modulation of excitability. *Nature* **366**, 322–324 (1993).
15. Welsh, B., Gomati, S. & Burgess, A. Three-dimensional chemical waves in the Belousov-Zhabotinsky reaction. *Nature* **304**, 611–614 (1983).
16. Jahnke, W., Henze, C. & Winfree, A. Chemical vortex dynamics in three-dimensional excitable media. *Nature* **336**, 662–665 (1988).
17. Panfilov, A. V., Rudenko, A. N. & Krinsky, V. I. Vortex rings in three-dimensional active media with diffusion in two components. *Biophysics* **31**, 926–931 (1986).
18. Biktahsev, V. N., Holden, A. V. & Zhang, H. Tension of organizing filaments of scroll waves. *Phil. Trans. R. Soc. Lond. A* **347**, 611–630 (1994).
19. Mironov, S., Vinson, M., Mulvey, S. & Pertsov, A. Destabilization of three-dimensional rotating chemical waves in an inhomogeneous BZ reaction. *J. Phys. Chem.* **100**, 1975–1983 (1996).
20. Pertsov, A. M., Aliev, R. R. & Krinsky, V. I. Three-dimensional twisted vortices in an excitable chemical medium. *Nature* **345**, 419–421 (1990).
21. Pertsov, A. M., Vinson, M. & Müller, S. C. Three-dimensional reconstruction of organizing centers in excitable chemical media. *Physica D* **63**, 233–235 (1993).
22. Keener, J. P. & Tyson, J. J. Helical and circular scroll wave filaments. *Physica D* **44**, 191–202 (1990).
23. Aliev, R. R. & Rovinsky, A. B. Spiral waves in the homogeneous and inhomogeneous Belousov-Zhabotinsky reaction. *J. Phys. Chem.* **96**, 732–736 (1992).
24. Keener, J. P. The dynamics of three-dimensional scroll waves in excitable media. *Physica D* **31**, 269–276 (1988).
25. Agladze, K. I., Kocharyan, R. A. & Krinsky, V. I. Direct observation of vortex ring collapse in a chemically active medium. *Physica D* **49**, 1–4 (1991).
26. Rudenko, A. N. & Panfilov, A. V. Drift and interaction of vortices in two-dimensional heterogeneous active medium. *Studia Biophysica* **98**, 183–188 (1983).
27. Pertsov, A. M., Davidenko, J. M., Salomonsz, R., Baxter, W. T. & Jalife, J. Spiral waves of excitation underlie reentrant activity in isolated cardiac muscle. *Circ. Res.* **72**, 631–650 (1993).

Acknowledgements. We thank M. Welner, J. Jalife, R. Gray and W. Baxter for reading the manuscript. This work was supported in part by the National Heart and Blood Institute, the Hendricks Fund, and the American Heart Association. Some of the numerical calculations were conducted using the computational resources of the Northeast Parallel Architectures Center (NPAC) at Syracuse University.

Correspondence should be addressed to A.P. (e-mail: pertsov@sunbird.pharm.hscsyr.edu).

Distribution and cycling of terrigenous dissolved organic matter in the ocean

Stephen Opsahl & Ronald Benner

Marine Science Institute, University of Texas at Austin, 750 Channelview Drive, Port Aransas, Texas 78373, USA

Terrigenous dissolved organic matter (DOM) is continuously discharged by rivers into the ocean, yet its distribution and reactivity within ocean basins remain poorly defined aspects of the global carbon cycle¹. Direct evidence for the presence of terrigenous DOM in the open ocean has been found only in dissolved humic substances isolated from surface waters of the eastern equatorial Pacific Ocean². Here we report the detection of low concentrations of lignin—a biopolymer found only in terrestrial vegetation—in DOM collected from the Pacific and Atlantic oceans, indicating that terrigenous DOM is distributed throughout the ocean water column. Higher lignin concentrations in Atlantic waters, relative to Pacific waters, reflect terrigenous DOM concentrations that are 2.6 times higher in the Atlantic. This finding is consistent with the 3.6-times greater riverine water-discharge to the Atlantic Ocean^{3,4}, and with known patterns of ocean circulation⁵. It appears that terrigenous DOM comprises only a small fraction (0.7–2.4%) of the total DOM in the ocean, and that its oceanic residence time (21–132 yr) is much shorter than that of marine DOM^{6,7}. The regeneration of nutrients during

rapid cycling of terrigenous DOM could contribute to high rates of primary production in the coastal ocean.

The annual riverine discharge ($\sim 0.4 \times 10^{15}$ g) of DOM and particulate organic matter to the oceans⁸ consists primarily of the degraded remains of terrestrial vegetation and is considered to be refractory to biodegradation^{9,10}. Terrigenous DOM has characteristically elevated C/N ratios, depleted stable carbon isotope ratios, and higher concentrations of aromatic carbon relative to marine DOM^{11,12}. However, bulk chemical parameters alone prove insufficient for quantifying the terrigenous component within the marine DOM reservoir because regional variability in sources and diagenetic alterations are not well understood. Unique biochemical tracers, such as lignin, offer much greater sensitivity for the quantification of terrigenous material in the ocean. The lignin biomarker approach is based on the identification and quantification of lignin-derived oxidation products which provide definitive evidence for the presence of vascular plant material, and thus carry an unambiguous terrigenous signature^{2,9}. Lignin oxidation products also provide information on the taxonomic source and diagenetic state of terrigenous organic matter^{9,13}.

Seawater samples were collected in 1992 from various depths and locations in the Pacific and Atlantic oceans (Table 1). Samples were passed through 0.2 or 0.1 μ m pore-size filters to remove particles, and the DOM was concentrated for subsequent lignin analysis using tangential-flow ultrafiltration with polysulphone filters^{12,14}. A substantial fraction (14–31%) of the DOM in seawater samples was recovered as a dry powder. Using the same isolation techniques, DOM was also collected from the Amazon and Mississippi rivers¹⁵. These rivers represent both tropical and temperate drainage basins and their combined flows account for $\sim 20\%$ of global riverine discharge³. Aeolian transport of particulate terrigenous organic matter to the ocean is quantitatively significant¹⁶, but the contribution of atmospheric deposition to terrigenous DOM in the ocean is currently unknown.

Ultrafiltered DOM was subjected to alkaline CuO oxidation¹⁷ to release lignin-derived phenols. The six vanillyl and syringyl phenols

Table 1 Properties of DOM collected from the Pacific and Atlantic oceans

Latitude	Longitude	Depth (m)	DOC (μ M)	DOC recovered (%)	$\delta^{13}\text{C}^*$ (‰)
Pacific Ocean					
18°47' N	133°50' W	2	78	24	–21.7
01°30' N	140°00' W	2	72	31	–21.6
12°12' S	134°40' W	2	82	27	–21.4
02°00' S	140°00' W	100	68	22	–21.8
11°58' S	135°07' W	100	85	21	–21.7
12°06' S	134°55' W	200	57	22	–21.5
12°19' S	134°26' W	375	53	20	–21.7
02°00' S	140°00' W	400	56	16	–21.7
02°00' S	140°02' W	400	51	23	–21.6
01°57' S	140°03' W	4,000	44	18	–21.8
12°00' S	135°00' W	4,000	45	14	–21.6
Atlantic Ocean					
31°50' N	64°10' W	1	72	23	–22.2
31°50' N	64°10' W	900	47	17	–21.7
31°50' N	64°10' W	2,400	46	23	–22.2

Shown here are sample locations, sampling depths, dissolved organic carbon (DOC) concentrations, DOC recoveries, and stable carbon isotope ratios ($\delta^{13}\text{C}$) of ultrafiltered dissolved organic matter (DOM). Amicon DC30 and DC10L ultrafiltration systems with 1,000 molecular weight cut-off polysulphone filters were used for the isolation of DOM^{12,14}. The ultrafiltration and CuO oxidation equipment and procedures were tested as sources of lignin contamination by isolating DOM from 2001 phytoplankton cultures grown in artificial sea water. Very low concentrations of lignin phenols (<5% of open-ocean DOM) were detected, indicating the isolation and analysis procedures were not a significant source of lignin contamination.

* Expressed relative to the PDB standard.

(vanillin, acetovanillone, vanillic acid, syringaldehyde, acetosyringone and syringic acid) are the most abundant lignin-derived oxidation products. Of these, only acetosyringone was below the detection limit and could not be positively identified. The two cinnamyl phenols, *p*-coumaric acid and ferulic acid, are often reported among the suite of lignin oxidation products, but they were not positively identified in these marine DOM samples.

Direct measurements of the vertical distribution of dissolved lignin phenols in different ocean basins provides new information regarding the distribution and reactivity of lignin and the associated terrigenous component of DOM. Measurable levels of lignin in all water masses sampled reveal for the first time the presence of dissolved terrigenous organic matter in the deep oceans (Fig. 1a). Open-ocean lignin concentrations are lower than values reported for coastal waters¹⁸ and are fairly uniformly distributed with depth. Large differences in lignin phenol concentrations between the Atlantic (24.9–28.5 ng l⁻¹) and Pacific (6.0–14.2 ng l⁻¹) ocean appear to reflect the disproportionately large share of global riverine discharge into the Atlantic (55%) relative to the Pacific (31%), which amounts to a 3.6-fold higher input to the Atlantic on a per volume basis^{3,4}. The observation of higher concentrations of terrigenous DOM in the Atlantic relative to the Pacific is consistent with known patterns of global ocean circulation⁵ and indicates a much greater influence of continental processes on biogeochemical cycling in the Atlantic relative to the Pacific ocean.

Stable carbon isotope compositions ($\delta^{13}\text{C}$) of all ultrafiltered DOM samples fall within the narrow range of -22.2‰ to -21.4‰ (Table 1) and are similar to values reported for marine DOC^{6,7}. These values are indicative of a marine phytoplankton source, but carbon isotopic compositions alone are insufficient to resolve a minor terrigenous component of marine DOM. Carbon-normalized yields of lignin-derived phenols (A_6) from rivers and major ocean basins can be used to quantitatively estimate the terrigenous component of marine DOM. Assuming that ultrafiltration recovers a representative fraction of dissolved lignin and carbon from ocean and river endmembers, we estimate that ~0.7% and 2.4% of the DOM in the Pacific and Atlantic, respectively, is of terrigenous origin (Table 2). Importantly, the lignin phenol data reveal a large (three-fold) difference between the fraction of terri-

genous DOM in the Atlantic and Pacific oceans which is not evident from stable carbon isotope measurements.

The distribution and abundance of individual lignin phenols provides a compositional fingerprint which often yields complementary information regarding the source and extent of diagenetic alteration of naturally occurring terrigenous materials. Syringyl phenols, unique oxidation products of angiosperm lignins, have not been previously identified in open-ocean DOM. The presence of syringyl phenols in all DOM samples from this study demonstrates that angiosperms are major contributors to the terrigenous component of oceanic DOM. The abundance of syringyl relative to vanillyl phenols (S/V) is indicative of the amount of angiosperm tissue in a given source¹³. The S/V of Amazon and Mississippi river DOM was in the range 0.5–1.2, reflecting the predominance of angiosperms in these tropical and temperate drainage basins. The S/V values in seawater DOM samples were all consistently lower than those in river samples (Fig. 1b). The relatively low S/V (0.1–0.4) in seawater DOM could reflect terrigenous contributions from geographical regions dominated by gymnosperms, such as high-latitude rivers. However, selective degradation of syringyl relative to vanillyl phenols can also explain the observed trends. Additional studies are required to resolve these possibilities. If smaller-scale regional differences in S/V are observed, then the potential exists to distinguish contributions from rivers draining catchment basins

Table 2 Occurrence and reactivity of terrigenous DOM in the ocean

	A_6	Terrigenous DOM* (%)	Lignin phenols (ng l ⁻¹)	Residence time† (yr)
Pacific	6.6	0.7	10.7	29–132
Atlantic	21.3	2.4	27.3	21–93

Shown here are carbon-normalized lignin yields (A_6 ; μg per 100 mg of organic carbon), the estimated percentage of terrigenous organic matter in seawater dissolved organic matter (DOM), lignin phenol concentrations and average residence times for dissolved terrigenous organic matter in the Pacific and Atlantic oceans.

* Given by $(A_6 \text{ ocean}/A_6 \text{ river}) \times 100$, where A_6 river is the average value (895 μg per 100 mg organic carbon) for Amazon and Mississippi river DOM.

† Calculated using river discharge and ocean volumes provided in refs 3 and 4, and an average lignin phenol concentration of 21.6 $\mu\text{g l}^{-1}$ for the Amazon and Mississippi rivers. Shorter residence times were calculated using measured concentrations of lignin phenols. Longer residence times were calculated assuming ultrafiltration recovered a similar fraction (22%) of lignin phenols and dissolved organic carbon from sea water.

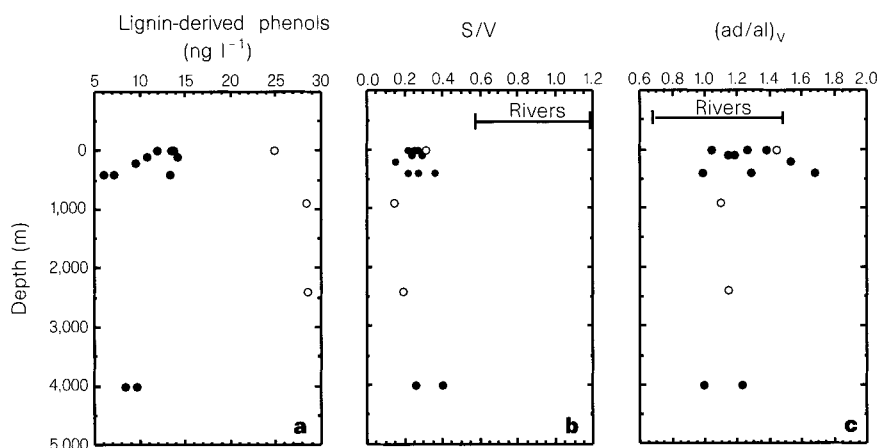


Figure 1 Depth profiles of lignin parameters in ultrafiltered dissolved organic matter from the Pacific (closed circles) and Atlantic (open circles) oceans. Each value represents a single determination from individual water samples collected at the depths indicated. **a**, Lignin-derived phenol concentrations; **b**, ratios of syringyl (syringaldehyde + acetosyringone + syringic acid) to vanillyl (vanillin + acetovanillone + vanillic acid) phenols; and **c**, ratios of vanillic acid to vanillin. 'Rivers' depicts the range of measured values ($n = 5$) in samples of Amazon and Mississippi river DOM. To measure lignin-derived phenols, DOM samples were oxidized in alkaline CuO at 155 °C for 3 h (ref. 17). Samples were acidified to pH 1

with HCl, extracted in ethyl ether, and dried under N_2 . Subsamples were resuspended in pyridine and converted to trimethylsilyl derivatives before analysis. The resulting mixture was separated using a Hewlett Packard 5890A gas chromatograph equipped with an HP-5MS capillary column (30 m, 0.25 mm internal diameter). Positive identification was achieved using a Hewlett Packard 5972 mass-selective detector, based on comparison with commercial standards. Quantification was accomplished using selected ion monitoring and ethylvanillin as the internal recovery standard.

dominated by angiosperm or gymnosperm vegetation. If syringyl phenols are selectively removed during the decomposition of terrigenous DOM in the ocean then low S/V could be indicative of specific diagenetic processes, such as microbial and photochemical oxidation.

During the microbial degradation of vascular plant tissues, ratios of vanillic acid to the aldehyde vanillin, (ad/al)_v, increase reflecting the extent of oxidative degradation of lignin¹⁹. The (ad/al)_v in oceanic DOM samples ranged from 0.9 to 1.7 (Fig. 1c) and did not demonstrate any obvious differences with depth. These values are similar to those reported for dissolved humic substances from surface Pacific waters² and are indicative of highly oxidized DOM. In this study, (ad/al)_v averaged 1.2 in all ocean DOM samples and was close to the average (1.1) for Amazon and Mississippi river DOM. Thus, although most of the terrigenous organic matter discharged into oceans is degraded, large changes in (ad/al)_v were not evident. These observations indicate that the mechanism(s) responsible for the removal of terrigenous DOM from the ocean does not alter its (ad/al)_v.

Average residence times for terrigenous organic matter in the Pacific and the Atlantic fall within the range 21–132 yr (Table 2), assuming dissolved lignin concentrations in the Amazon and Mississippi rivers are representative of global riverine discharge. Such short residence times demonstrate that most terrigenous organic matter cycles much more rapidly than the mixing time (500–1,000 yr) of the ocean^{5,20}. In comparison, deep-ocean DOC concentrations (~40 µM) are less variable between ocean basins than dissolved lignin concentrations, and bulk DOC has^{6,7} an average apparent age of 4,000–6,000 yr. To maintain the observed concentration gradient of terrigenous organic matter between the Pacific and Atlantic oceans, the dissolved terrigenous component must survive fewer ocean mixing cycles and turn over more rapidly than the bulk DOC reservoir. Notably, the terrigenous component (0.3 µM C) observed in the deep Pacific Ocean, the oldest water sampled in this study, appears to cycle on a timescale >150 yr, demonstrating the persistence of a longer-lived terrigenous component.

Some terrigenous organic matter is removed from solution during the mixing of fresh and salt waters²¹, but riverine transport of most terrigenous DOC²² and dissolved organic nitrogen²³ to the ocean is conservative, suggesting it is not reactive in the coastal ocean. In contrast, the short average residence time of most terrigenous DOM in the ocean suggests that removal mechanisms are relatively rapid and likely active in ocean margin regions. Rapid decreases in concentrations of lignin phenols in Atlantic shelf waters¹⁸ indicate that some terrigenous DOM is removed in the coastal zone. Reported rates of both photochemical^{24–26} and microbial²⁷ oxidation of riverine DOM are sufficient to oxidize most terrigenous DOM on timescales of months to years, which is similar to the residence times of water on continental shelves. Both of these processes are active in coastal waters and they regenerate inorganic nutrients^{28,29} that often limit marine primary production. These observations indicate that terrigenous DOM may be a quantitatively important source of nutrients to the coastal ocean, which accounts for ~25% of global ocean productivity³⁰. □

Received 29 July 1996; accepted 19 February 1997.

- Hedges, J. I. Global biogeochemical cycles: progress and problems. *Mar. Chem.* **39**, 67–93 (1992).
- Meyers-Schulte, K. J. & Hedges, J. I. Molecular evidence for a terrestrial component of organic matter dissolved in ocean water. *Nature* **321**, 61–63 (1986).
- Baumgartner, A. & Reichel, E. *The World Water Balance* (Oldenbourg, Munich, 1975).
- Sverdrup, H. U., Johnson, M. W. & Fleming, R. H. *The Oceans* (Prentice-Hall, Englewood Cliffs, NJ, 1942).
- Broecker, W. S. & Peng, T.-H. *Tracers in the Sea* (Lamont-Doherty Geological Observatory, Columbia Univ., Palisades, New York, 1982).
- Williams, P. M. & Druffel, E. R. M. Radiocarbon in dissolved organic matter in the central North Pacific Ocean. *Nature* **330**, 246–248 (1987).
- Bauer, J. E., Williams, P. M. & Druffel, E. R. M. ¹⁴C activity of dissolved organic carbon fractions in the north-central Pacific and Sargasso Sea. *Nature* **357**, 667–670 (1992).
- Meybeck, M. Carbon, nitrogen, and phosphorus transport by world rivers. *Am. J. Sci.* **282**, 401–450 (1982).
- Ertel, J. R., Hedges, J. I., Devol, A. H., Richey, J. E. & Ribeiro, M. Dissolved humic substances in the Amazon River system. *Limnol. Oceanogr.* **31**, 739–754 (1986).

- Ittekkot, V. Global trends in the nature of organic matter in river suspensions. *Nature* **332**, 436–438 (1988).
- Hedges, J. I., Hatcher, P. G., Ertel, J. R. & Meyers-Schulte, K. J. A comparison of dissolved humic substances from seawater with Amazon River counterparts by ¹³C-NMR spectrometry. *Geochim. Cosmochim. Acta* **56**, 1753–1757 (1992).
- Benner, R., Pakulski, J. D., McCarthy, M., Hedges, J. I. & Hatcher, P. G. Bulk chemical characteristics of dissolved organic matter in the ocean. *Science* **255**, 1561–1564 (1992).
- Hedges, J. I. & Mann, D. C. The characterization of plant tissues by their lignin oxidation products. *Geochim. Cosmochim. Acta* **43**, 1803–1807 (1979).
- Benner, R. in *Marine Particles: Analysis and Characterization* (eds Hurde, D. C. & Spencer, D. W.) 181–185 (Geophys. Monogr. 63, American Geophysical Union, Washington DC, 1991).
- Benner, R. & Hedges, J. I. A test of the accuracy of freshwater DOC measurements by high-temperature catalytic oxidation and UV-promoted persulfate oxidation. *Mar. Chem.* **41**, 161–165 (1993).
- Zafriou, O. C., Gagosian, R. B., Peltzer, E. T. & Alford, J. B. Air-to-sea fluxes of lipids at Enewetak Atoll. *J. Geophys. Res.* **90**, 2409–2423 (1985).
- Hedges, J. I. & Ertel, J. R. Characterization of lignin by capillary chromatography of cupric oxide oxidation products. *Anal. Chem.* **54**, 174–178 (1982).
- Moran, M. A., Pomeroy, L. R., Sheppard, E. S., Atkinson, L. P. & Hodson, R. E. Distribution of terrestrially derived dissolved organic matter on the southeastern U.S. continental shelf. *Limnol. Oceanogr.* **36**, 1134–1149 (1991).
- Opsahl, S. & Benner, R. Early diagenesis of vascular plant tissues: lignin and cutin decomposition and biogeochemical implications. *Geochim. Cosmochim. Acta* **59**, 4889–4904 (1995).
- Stuiver, M., Quay, P. D. & Ostlund, H. G. Abyssal water carbon-14 distribution and the age of the world oceans. *Science* **219**, 849–851 (1983).
- Sholkovitz, E. R. The removal of dissolved humic acids and iron during estuarine mixing. *Earth Planet. Sci. Lett.* **40**, 130–136 (1978).
- Mantoura, R. F. C. & Woodward, F. M. S. Conservative behaviour of riverine dissolved organic carbon in the Severn Estuary: chemical and geochemical implications. *Geochim. Cosmochim. Acta* **47**, 1293–1309 (1983).
- Gardner, W. & Stephens, J. A. Stability and composition of terrestrially derived dissolved organic nitrogen in continental shelf surface waters. *Mar. Chem.* **6**, 335–342 (1978).
- Amon, R. M. W. & Benner, R. Photochemical and microbial consumption of dissolved organic carbon and dissolved oxygen in the Amazon River. *Geochim. Cosmochim. Acta* **60**, 1783–1792 (1996).
- Kieber, R. J., Zhou, X. & Mopper, K. Formation of carbonyl compounds from UV-induced photodegradation of humic substances in natural waters: fate of riverine carbon in the sea. *Limnol. Oceanogr.* **35**, 1503–1515 (1990).
- Miller, W. L. & Zepp, R. G. Photochemical production of dissolved inorganic carbon from terrestrial organic matter: significance to the oceanic organic carbon cycle. *Geophys. Res. Lett.* **22**, 417–420 (1995).
- Benner, R., Opsahl, S., Chin-Leo, G., Richey, J. E. & Forsberg, B. R. Bacterial carbon metabolism in the Amazon River. *Limnol. Oceanogr.* **40**, 1262–1270 (1995).
- Pakulski, J. D., Benner, R., Amon, R. M. W., Eadie, B. & Whitedge, T. Community metabolism and nutrient cycling in the Mississippi River plume: evidence for intense nitrification at intermediate salinities. *Mar. Ecol. Prog. Ser.* **117**, 207–218 (1995).
- Bushaw, K. L. *et al.* Photochemical release of biologically available nitrogen from aquatic dissolved organic matter. *Nature* **381**, 404–407 (1996).
- Martin, J. H., Knauer, G. A., Karl, D. M. & Broenkow, W. W. VERTEX: carbon cycling in the northeast Pacific. *Deep-Sea Res.* **34**, 267–285 (1987).

Acknowledgements. We thank P. Quay for discussions about ocean mixing processes and M. McCarthy for assistance with the collection of samples. We also thank A. Skoog and the UTMSI biogeochemistry group for reviewing the manuscript. This work was supported by the US National Science Foundation.

Correspondence and requests for materials should be addressed to R.B. (e-mail: benner@utmsi.zo.utexas.edu).

Effect of mechanical interactions on the scaling of fracture length and aperture

C. E. Renshaw & J. C. Park

Department of Geology, State University of New York at Buffalo, Buffalo, New York 14260, USA

The aperture (or opening) of a fracture indicates the energy available for fracture growth and controls fracture permeability. The relationship between aperture and fracture length can therefore be used to infer the factors affecting fracture formation at different length scales and is of practical importance to hydrogeologists and petroleum engineers. A recent study¹ of the scaling properties of tensile fractures in the Krafla fissure swarm, Iceland, revealed a distinct break in slope in the aperture–length scaling relationship, corresponding to fractures a few metres in length: this break in slope was interpreted qualitatively as indicative of non-universal, scale-dependent growth mechanisms¹. Here we show, using quantitative fracture simulations, that the observed non-universal scaling of fracture apertures can be reproduced without recourse to multiple growth mechanisms. We argue that

the break in slope is instead intrinsic to the fracturing process and represents the maximum length scale at which the apertures of smaller fractures are affected by stress perturbations induced by larger fractures.

Linear elastic fracture mechanics (LEFM) theory has been successfully applied to many aspects of rock fracture², including fault displacement³ and dyke aperture distributions⁴. When integrated with numerical methods for stress analysis, LEFM models can accurately simulate the propagation of experimentally induced fractures^{5,6}.

LEFM theory predicts that the maximum aperture w of an isolated mode I fracture scales linearly with fracture length $2a$ (ref. 7):

$$w = \sigma \frac{(1 - \nu)}{\mu} (2a) \quad (1)$$

where σ is the effective driving stress (remote tension plus the internal fluid pressure) and ν and μ are Poisson's ratio and the shear modulus, respectively. When multiple fractures are present, the stress perturbations induced by each fracture will interact, in some locations positively and in others negatively. This interaction affects the growth of the fractures and the distribution of their apertures. For example, interaction of *en echelon* segments making up larger fractures results in sublinear aperture-length scaling⁸. It is shown here that mechanical interaction can also result in superlinear aperture-length scaling.

Previous work has demonstrated that the growth and interaction of multiple fractures can be accurately modelled⁶. In this model, small, equal-length, parallel flaws are randomly distributed within a two-dimensional linear elastic medium. Analytical solutions for the stress field around an isolated fracture are used to construct a matrix of influence coefficients representing the mechanical interaction between each pair of fractures. This matrix is used to calculate the average traction along each fracture⁹ and the energy available for the growth of each fracture tip, quantified as the energy release rate G (ref. 10). The maximum aperture along each fracture is estimated from the average tractions and energy release rates using the method of polynomially distorted ellipses⁹. To simulate fracture growth, each tip is advanced a distance proportional to the available energy and the average tractions recalculated. Model results compare

favourably to the growth observed in experimentally generated fracture networks⁶. More detailed discussions of the numerical methods are contained in references 6 and 9.

A key aspect of the fracture model is the relationship between G and the fracture growth rate v_f . For subcritical fracture propagation, this relationship can be described by an equation of the form

$$v_f \propto G^\alpha \quad (2)$$

where α is a constant referred to as the growth rate exponent. For a constant rate of stress increase, $\alpha = 1$ defines a boundary between stable or decelerating ($\alpha < 1$) and unstable or accelerating ($\alpha > 1$) fracture growth¹¹. Experimental analyses of single fractures¹² suggest that when fracture propagation is controlled by the transport of reactive species to the fracture tip or by the coalescence of cavities ahead of the fracture tip, α is in the range 1–5. Alternatively, when fracture growth is limited by the kinetics of the chemically induced weakening of the silicate bond at the tip of the fracture, α is in the range 10–25. For multiple fractures, inversions of acoustic emission data suggest relatively low values of α (0.35–1.75), depending on the confining pressure¹³. More recently Liakopoulou *et al.*¹⁴ have shown a sudden change from stable ($\alpha = 0.7$) to unstable ($\alpha \approx 2.5$) fracture growth during fault nucleation by coalescence of tensile microfractures. Although the acoustic inversion results were obtained in compression tests, a similarly low value of α (0.35) was found for the growth of multiple fractures in brittle coating under tension⁶. It is unclear why values of α measured for single fractures are significantly greater than those inferred from the growth of multiple fractures.

In using the fracture model to simulate the Krafla fracture swarm, it is assumed that the Krafla fractures are primarily opening mode, as is likely in active rift zones. Application of the model then requires information on the appropriate values for the growth rate exponent, the flaw density and the fracture density, which is a measure of the total inelastic strain. Field observations indicate that the Krafla fractures initiated from columnar cooling joints that have a characteristic dimension of about 0.3 m (ref. 1). These joints are arranged hexagonally, suggesting a uniform distribution of abutting flaw orientations. However, under predominantly uniaxial stress states, such as at Krafla, fractures will preferentially initiate from flaws oriented normal to the direction of greatest tensile stress, so we

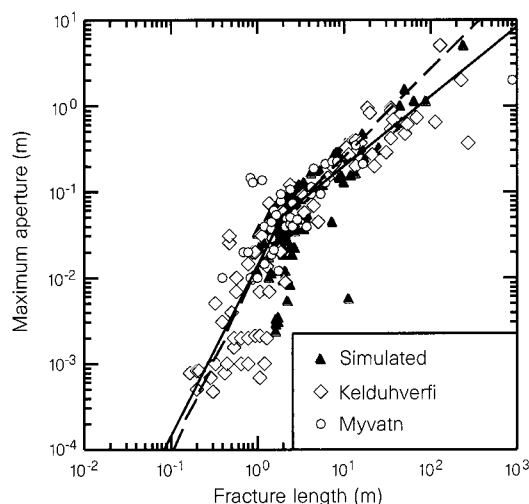


Figure 1 Scaling of fracture aperture with length at two locations within the Krafla fracture swarm and simulated scaling using $\alpha = 2.5$. Solid and dashed lines represent least-squares best fits of curves of the form given in equation (4) to the measured and simulated data, respectively. Krafla data after Hatton *et al.*¹. Simulated apertures are determined assuming $\sigma(1 - \nu)/\mu = 10^{-2}$.

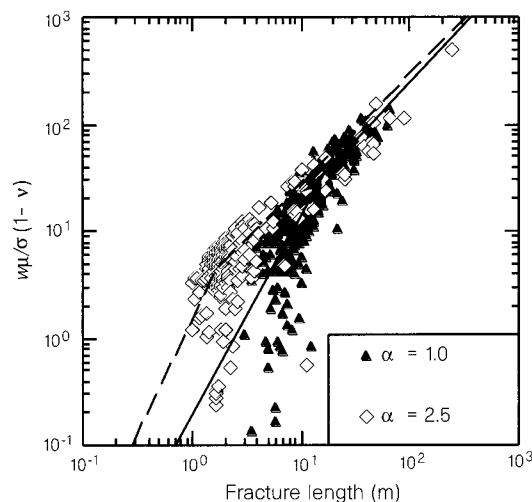


Figure 2 Scaling of fracture aperture with length in the simulated fracture set for two values of α . Apertures are normalized by the shear modulus μ , the driving stress σ , and Poisson's ratio ν . Lines indicate best fit curves for $\alpha = 1.0$ (solid) and $\alpha = 2.5$ (dashed).

consider only randomly distributed flaws with this orientation.

A minimum estimate of the flaw density D_f can be determined as:

$$D_f = \frac{na^2}{A} \quad (3)$$

where a is half the characteristic dimension of the cooling joints and n is the number of fractures measured within a mapped area of size A . This represents a minimum flaw density as not all fractures within the given area are mapped. In the Kelduhverfi area in the north of the fracture zone, Hatton *et al.*¹ mapped 80 fractures in an area of 91,000 m², implying $D_f \approx 2 \times 10^{-5}$. Near Myvatn at the south end of the fracture swarm, the authors mapped roughly 40 fractures in an area of 18,000 m², implying $D_f \approx 5 \times 10^{-5}$. Using these values as a guide, a flaw density of 1×10^{-4} is used for the results presented here. Previous work demonstrates that model results are relatively insensitive to flaw density^{6,15}.

The average fracture density D_s of the Krafla fracture swarm can be determined using equation (3), with D_s replacing D_f and a representing the average half-length of the fractures. However, local measurements of fracture density, such as derived from the work of Hatton *et al.*¹, can be inaccurate because fractures tend to cluster¹⁶. From field work and detailed studies of aerial photographs, Opheim and Gudmundsson¹⁷ identified 1,083 fractures within an area of 7.9×10^7 m² of the Krafla fracture zone (more than 700 times the area considered by Hatton *et al.*¹). The average half-length of these fractures is 17 m, resulting in a fracture density of 0.4. This density is within the range of typical densities determined from published fracture trace maps of surficial outcrops¹⁵.

As discussed above, the appropriate value of α is uncertain, but likely to be in the range 0.35–25. Experimental data indicate that appropriate values of α for the growth of multiple fractures are typically in the lower part of this range ($\alpha < 3$). Given these results, multiple simulations with values of α in the range 0.5–10 were performed to investigate the influence of α on aperture scaling.

The scaling of simulated fracture apertures for two values of α is shown in Fig. 2. As observed in the field, there is a distinct break in slope for both values of α . To quantify these results, a multi-scale log–linear curve of the form

$$\begin{aligned} W &= s_1(L - L_0) + W_0 & L \leq L_0 \\ W &= s_2(L - L_0) + W_0 & L > L_0 \end{aligned} \quad (4)$$

was fitted to the data using a least-squares method, where W is the log-maximum aperture, L the log-fracture length, and s_1 and s_2 the scaling exponents (Table 1). Subscripts indicate the fracture length and aperture corresponding to the break in slope. Fits were obtained using Powell's method¹⁸.

Below L_0 , apertures scale superlinearly ($s_1 > 1$) with fracture length, whereas above the break the scaling is approximately linear ($s_2 \approx 1$). Because the simulated fractures are not segmented,

the sublinear scaling of the larger fractures in the Krafla swarm is not reproduced. However, the superlinear scaling of the smaller fractures, the increased scatter in the scaling of the smaller fractures and the distinct break in slope are all consistent with the Krafla data. Different values of α do not significantly change the form of the aperture scaling, but do influence the fracture length at which the break in slope occurs. Table 1 suggests $\alpha \approx 2$ for the Krafla fractures. The simulated aperture–length scaling for $\alpha = 2.5$ is compared to the Krafla data in Fig. 1. Notice that the break in slope occurs even though the mechanism of fracture growth, controlled by the value of α , is constant. Hence the model reproduces the observed non-universal scaling of the fracture apertures without recourse to multiple growth mechanisms.

The superlinear aperture–length scaling of the smaller fractures suggests that they are affected by mechanical interaction. As the aperture of a fracture is an indication of the energy available for fracture growth (for example, for isolated fractures $G \propto a$; ref. 10), superlinear scaling indicates that smaller fractures have less energy available for growth than they would if they were isolated. This implies that as the density of a fracture set increases, mechanical interaction inhibits the growth of smaller fractures, localizing fracture growth to the longest fractures in the system. Despite this inhibition, it can be seen in Fig. 1 that the smallest fractures in the simulated set are slightly longer than observed by Hatton *et al.*¹. This is a consequence of the model assumption that all the flaws are of equal length and that there is no blunting of the flaw tips. All the flaws therefore begin to grow at the same time and at the same rate. In reality, longer flaws with sharper tips grow first, followed later by the smaller flaws and flaws with tips that are blunt. These later flaws will not grow much, resulting in the numerous smaller fractures in the Krafla field data.

Our results highlight the dangers of applying universal scaling laws, even when the deformational mechanisms, such as fracture propagation, are universal. ■

Received 16 September 1996; accepted 11 February 1997.

- Hatton, C. G., Main, I. G. & Meredith, P. G. Non-universal scaling of fracture length and opening displacement. *Nature* **367**, 160–162 (1994).
- Pollard, D. D. & Aydin, A. Progress in understanding jointing over the past century. *Geol. Soc. Am. Bull.* **100**, 1181–1204 (1988).
- Bürgmann, R., Pollard, D. D. & Martel, S. J. Slip distributions on faults: effects of stress gradients, inelastic deformation, heterogeneous host-rock stiffness and fault interaction. *J. Struct. Geol.* **14**, 1133–1148 (1994).
- Delaney, P. T. & Pollard, D. D. Deformation of host rocks and flow of magma during growth of minette dikes and breccia-bearing intrusions near Ship Rock, New Mexico. *US Geol. Surv. Prof. Paper* 1350 (1981).
- Thomas, A. L. & Pollard, D. D. The geometry of echelon fractures in rock: Implications from laboratory and numerical experiments. *J. Struct. Geol.* **15**, 323–334 (1993).
- Renshaw, C. E. & Pollard, D. D. Numerical simulation of fracture set formation: a fracture mechanics model consistent with experimental observations. *J. Geophys. Res.* **99**, 9359–9372 (1994).
- Pollard, D. D. & Segall, P. in *Fracture Mechanics of Rock* (ed. Atkinson, B. K.) 277–349 (Academic, San Diego, California, 1987).
- Vermilye, J. M. & Scholz, C. H. Relation between vein length and aperture. *J. Struct. Geol.* **17**, 423–434 (1995).
- Kachanov, M. Elastic solids with many cracks: A simple method of analysis. *Int. J. Solids Structures* **23**, 23–43 (1987).
- Broek, D. *Elementary Engineering Fracture Mechanics*, 1–516 (Martinus Nijhoff, Boston, 1986).
- Main, I. Prediction of failure times in the Earth for a time-varying stress. *Geophys. J.* **92**, 455–464 (1988).
- Atkinson, B. K. & Meredith, P. G. in *Fracture Mechanics of Rock* (ed. Atkinson, B. K.) 111–166 (Academic, New York, 1987).
- Main, I. G., Sammonds, P. R. & Meredith, P. G. Application of a modified Griffith criterion to the evolution of fractal damage during compressional rock failure. *Geophys. J. Int.* **115**, 367–380 (1993).
- Liakopoulou-Morris, F., Main, I. G. & Crawford, B. R. Microseismic properties of homogenous sandstone during fault nucleation and frictional sliding. *Geophys. J. Int.* **119**, 219–230 (1994).
- Renshaw, C. E. Influence of sub-critical fracture growth on the connectivity of fracture networks. *Wat. Resour. Res.* **32**, 1519–1530 (1996).
- Wu, H. & Pollard, D. D. An experimental study of the relationship between joint spacing and layer thickness. *J. Struct. Geol.* **17**, 887–905 (1995).
- Opheim, J. A. & Gudmundsson, A. Formation and geometry of fractures, and related volcanism, of the Krafla fissure swarm, northeast Iceland. *Geol. Soc. Am. Bull.* **101**, 1608–1622 (1989).
- Press, W. H., Teukolsky, S. A., Vetterling, W. T. & Flannery, B. P. *Numerical Recipes in C, The Art of Scientific Computing* 1–994 (Cambridge Univ. Press, New York, 1992).

Acknowledgements. We thank I. Main for comments on the manuscript.

Correspondence should be addressed to C.E.R. (e-mail: renschaw@acsu.buffalo.edu).

Table 1 Least-squares best fit parameters for measured and simulated fracture data

	L_0 (log-metre)	s_1	s_2
Krafla	0.32	1.98	0.83
$\alpha = 0.05$	1.26	2.54	1.31
$\alpha = 1.0$	1.21	1.89	1.10
$\alpha = 1.5$	0.74	2.46	1.28
$\alpha = 2.5$	0.24	2.13	1.02
$\alpha = 4.0$	0.15	1.76	0.70
$\alpha = 10$	–0.17	2.28	0.75

The evolution of apical dominance in maize

John Doebley*, Adrian Stec* & Lauren Hubbard†

* Department of Plant Biology, University of Minnesota, St Paul, Minnesota 55108, USA

† USDA Plant Gene Expression Center, Albany, California 94710 and Department of Plant Biology, University of California, Berkeley, California 94720, USA

The domestication of crop plants has often involved an increase in apical dominance (the concentration of resources in the main stem of the plant and a corresponding suppression of axillary branches)¹. A striking example of this phenomenon is seen in maize (*Zea mays* spp. *mays*), which exhibits a profound increase in apical dominance compared with its probable wild ancestor, teosinte (*Zea mays* ssp. *parviglumis*)². Previous research has identified the *teosinte branched1* (*tb1*) gene as a major contributor to this evolutionary change in maize³. We have cloned *tb1* by transposon tagging and show here that it encodes a protein with homology to the *cycloidea* gene of snapdragon⁴. The pattern of *tb1* expression and the morphology of *tb1* mutant plants suggest that *tb1* acts both to repress the growth of axillary organs and to enable the formation of female inflorescences. The maize allele of *tb1* is expressed at twice the level of the teosinte allele, suggesting that gene regulatory changes underlie the evolutionary divergence of maize from teosinte.

Teosinte plants typically bear an elongated lateral branch at most nodes on their main stems². These branches are tipped by male inflorescences (tassels) and the slender female inflorescences (ears) are borne on secondary branches in the axils of the leaves on the primary branches. In contrast, maize plants typically produce a lateral branch at only two or three of the nodes on their main stems, and these are short and tipped by ears. Previously, we demonstrated that these differences in plant architecture are governed by a small number of quantitative trait loci (QTL)⁵. One of these QTL was shown by genetic complementation testing to correspond to the *teosinte branched1* (*tb1*) locus, and it is this QTL that largely controls these differences in plant architecture³. In fact, the effects of *tb1* alone are sufficiently large that, when the maize chromosome segment carrying this gene is transferred into teosinte, it fully converts teosinte to maize plant architecture.

The *tb1* mutant of maize causes a complete loss of apical dominance, allowing the unrestrained outgrowth of axillary buds^{3,6}. At the basal nodes of the main stem, plants homozygous for the reference allele (*tb1-ref*) produce a profusion of tillers (basal branches) where wild-type plants typically produce few, if any, tillers because the outgrowth of most of their axillary buds is completely arrested (Fig. 1a, c). At some upper nodes, mutant plants bear long lateral branches tipped by tassels where wild-type maize plants have only short branches tipped by ears (Fig. 1b, d). Because both tillers and upper lateral branches arise from axillary meristems, in a general sense, *tb1* controls the fate of axillary meristems, resulting in different outcomes depending on the position of the meristem along the main culm. At lower nodes, *tb1* controls fully arrested buds versus tillers, and at upper nodes, it controls ear shoots versus elongated branches.

To characterize further the mutant phenotype, we studied a population of 135 F₂ plants segregating for *tb1-ref* in the background of the maize inbred line A158 (Table 1). The conversion of the inflorescences terminating the lateral branches from female (wild-type) to male (mutant) was fully recessive. However, the effect of *tb1-ref* on the number of tillers was semidominant, which is consistent with a previous report⁷. Semidominance was also shown

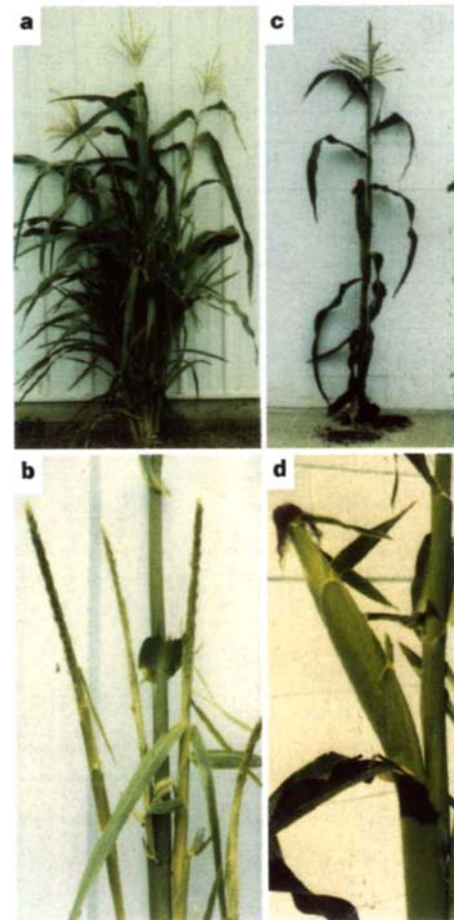


Figure 1 Plants and inflorescences of wild-type and *tb1-ref* mutant maize. **a, b**, *teosinte branched1* mutant. **c, d**, Wild-type maize (inbred A158).

in our population for the degree of tiller growth, the number of ear shoots and lateral branch length.

Comparison of mutant and wild-type plants shows that *tb1* increases the number of primary lateral branches and tillers on the plants by allowing the outgrowth of buds rather than by increasing the number of axillary buds formed on the main stalk. First, *tb1* does not affect the total number of nodes (each potentially bearing a single axillary bud) in the main culm (Table 1). Second, *tb1* does not extend branching up the main culm, that is, the node (position) of the uppermost lateral branch in *tb1* mutants is the same as that in wild-type plants. Third, the extra tillers in *tb1* mutants result from the primary tillers each having one or two secondary tillers.

tb1 also affects the development of secondary lateral branches. In teosinte, these branches bear ears; in modern maize, these branches usually do not develop. *tb1-ref* plants typically form sterile, tassellike inflorescences on many secondary lateral branches, although rarely they can produce small and poorly formed ears⁸. The fact that *tb1-ref* plants, unlike teosinte, do not normally produce ears on secondary branches highlights an important point. *tb1-ref* mutant plants have a teosinte-like phenotype only in their primary lateral branches and not in the secondary ones. Thus, *tb1* is not a truly atavistic mutation.

To isolate the *tb1* gene, we used the *Mutator* (*Mu*) transposable element system⁹. Homozygous *tb1-ref* plants were crossed to an active *Mu* stock, and 26,000 F₁ plants were grown. Among these, three new *tb1* mutants (*tb1-mum1*, *tb1-mum2* and *tb1-mum3*) were observed. Each new mutant was crossed to maize inbred A632, and

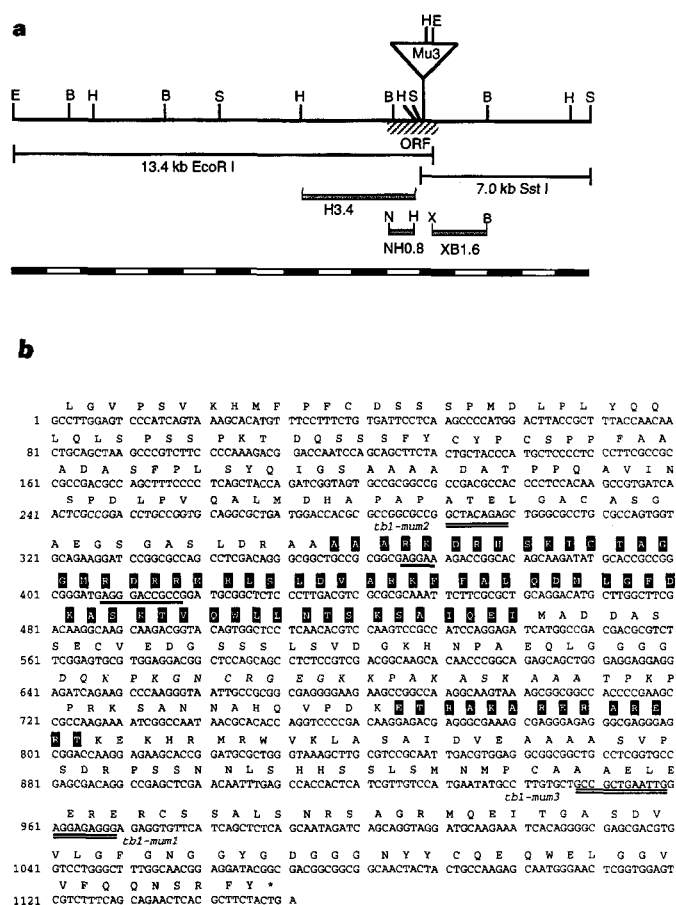


Figure 2 Structure of *teosinte branched1*. **a**, Restriction endonuclease maps of overlapping genomic λ clones (E13.4 and S7.0) of *tb1*. Insertion site of the *Mu3* element in the *tb1* open reading frame (ORF, hatched box) is shown. Three probes (H3.4, NH0.8 and XB1.6) used in Southern and northern blot analyses are shown. Abbreviations for restriction enzymes: B, *Bam*HI; E, *Eco*RI; H, *Hind*III; N, *Nco*I; S, *Sst*I; and X, *Xba*I. Not all N and X sites were mapped. Scale in 1 kb segments. **b**, Nucleic-acid and deduced amino-acid sequence of the *tb1* cDNA clone. Two domains conserved between *tb1*, *cycloidea* and *Arabidopsis* ESTs are highlighted. Putative bipartite nuclear localization signal is underlined. Sequences duplicated by the *Mu* insertions are double underlined.

Table 1 Morphometric comparison of mutant and wild-type alleles of *tb1*

Trait	Genotypes		
	<i>tb1/tb1</i>	<i>Tb1/tb1</i>	<i>Tb1/Tb1</i>
Tiller number per plant	8.00	1.41	0.24
Tiller growth as % of length of main stem	—	45.00	7.00
Number of visible axillary shoots	10.50	3.59	1.93
Length of uppermost lateral branch (cm)	52.51	24.06	16.21
Number of ear shoots	—	2.25	1.69
Node of the uppermost axillary shoot	10.85	10.85	10.52
Number of nodes on main stem	17.13	16.85	16.31
Number of plants measured	40	66	29

Bold underlining indicates the results of Scheffe F-tests such that a continuous line under adjacent values in a row indicates statistically equivalent values and a broken line between two values indicates statistically different values ($P = 0.05$). Plants were genotyped by Southern blot analysis with the H3.4 probe (Fig. 2a).

14 progeny from each of these crosses were grown to small seedlings and then used for DNA extraction. Southern blot analysis using marker loci that closely flank *tb1* was used to discriminate progeny that possessed the *tb1-ref* versus the new *Mu* alleles¹⁰. We also screened the progeny by Southern blot analysis with *Mu* element probes and thereby identified a *Mu3* element that cosegregated with *tb1-mum1*. We examined an additional 156 progeny from this family and found that this *Mu3* element was associated with the new mutant allele in 169 of the 170 plants. We attribute the single exception to either pollen contamination or a germinal reversion.

Overlapping genomic restriction fragments carrying the cosegregating *Mu3* element were cloned into λ vectors (Fig. 2a). Subclones of these were used for Southern blot analyses of the 20 full sibs of each of the three new mutant plants, showing that each of these mutants has an insertion into this region of the genome that did not exist in their sibs and the progenitor stocks (Fig. 3a). These observations provide the crucial evidence that we have newly tagged *tb1*. We sequenced portions of the λ clones and designed oligonucleotide primers to use in conjunction with a *Mu*-specific primer to amplify (using polymerase chain reaction (PCR)) the insertion fragments for each of the three *Mu* alleles. DNA sequence analysis of these fragments enabled us to identify the *Mu* insertion point of each allele (Fig. 2b).

Northern blot analysis with RNA from various tissues and both mutant and wild-type plants revealed a 1.5-kb message in wild-type maize axillary inflorescence (ear) primordia, immature internodes below these primordia, and immature husks surrounding these primordia (Fig. 3b). The *tb1-ref* allele has a larger message (about 2.5 kb) that was expressed in the same tissue types as the wild-type allele. The *tb1-ref* message is less intense than that of the wild-type allele, suggesting either that it is expressed at a lower level or its

message is less stable. The *tb-mum1* allele has no apparent message, suggesting that the *Mu* insertion may interfere with message stability. Teosinte produces the same size message as maize, indicating that maize evolution did not involve the creation or loss of *tb1*.

We used the 0.8-kb *Nco*I–*Hind*III fragment to screen a cDNA library derived from immature ear. We obtained a single clone with a 1,306 bp insert, excluding the poly(A) tail. The sequence of this clone (Fig. 2b) agreed fully with the genomic clone of maize inbred A619, but differed from the W22 genomic clone by a few polymorphisms that represent allelic differences. The genomic and cDNA sequences are fully collinear without any evidence for introns. This clone would appear to be nearly full length. Further experiments are needed to map the 5' end of the gene.

We searched the GenBank database and found that *tb1* shares two short regions of homology with the *cycloidea* gene of snapdragon⁴, and three *Arabidopsis* expressed sequence tags (ESTs: R29994, R30409, T45419) (Fig. 2b). First, a 62-amino-acid region is found in all sequences and a second shorter region is shared with one *Arabidopsis* EST, *cycloidea* and *tb1*. The conservation of these domains between maize, snapdragon and *Arabidopsis* indicates that they are important to protein function. The 62-amino-acid region of homology contains a putative nuclear localization signal, which suggests that these proteins may play a role in the regulation of transcription⁴.

There are several parallels between *cycloidea* and *tb1*. First, both mutants affect axillary structures, either flowers (*cycloidea*) or branches (*tb1*). Second, both genes have been proposed to function as repressors of organ growth^{3,4}. Third, both genes affect the growth of petals and stamens, organs whose development is regulated by the B class of floral organ identity genes¹¹. *Cyc*⁺ arrests the growth of both the dorsal petal and stamen of the snapdragon flower⁴. In the

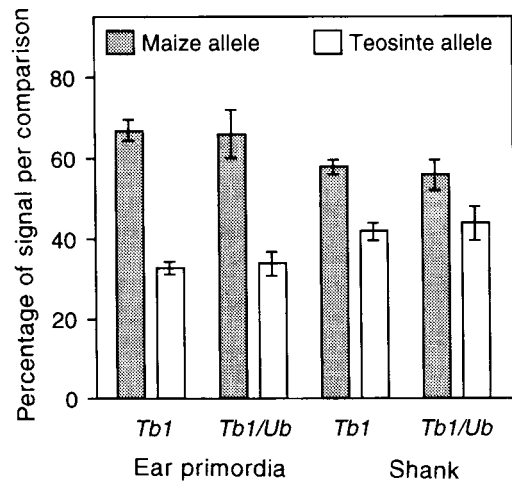
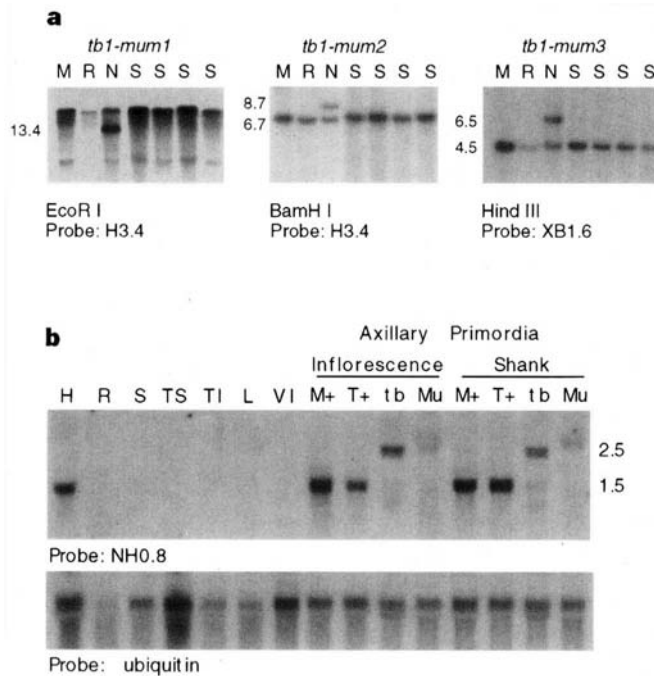


Figure 3 a, Southern blot analyses. *tb1-mum* alleles each contain an insertion in *tb1* as compared to the parental *tb1-ref* and *Mu* stocks, and 20 (4 shown) full sibs of each mutant plant. The *Mu3* element in the *tb1-mum1* allele possesses an *EcoRI* site, so this insertion created the 13.4 kb *EcoRI* fragment. The *Mu* insertions in the other two *tb1-mum* alleles increased the sizes of existing fragments by about 2 kb. Abbreviations: *Mu* stock, M; new *Mu* mutant plant from the tagging population, N; *tb1-ref* stock, R; full sibs of the new *Mu* mutant plant, S. **b**, Northern blot analyses. RNAs from inbred W22: ear husk, H; root, R; dark grown seedlings, S; immature tassel spikelets, TS; immature tassel rachis internodes, TI; immature leaves, L; and immature vegetative internodes of the main stem, VI. Additional RNAs from axillary inflorescence and shank primordia of plants homozygous for alleles: *Tb1 + Maize*, M+; *Tb1 + teosinte*, T+; *tb-ref*, tb; *tb1-mum1*, Mu.

Figure 4 *Tb1* message accumulation for ear primordia and immature shank of the maize (*Tb1 + Maize*) and teosinte (*Tb1 + teosinte*) alleles in inbred W22 genetic background. Data are presented both as the amount of *Tb1* message per 10 µg of total cellular RNA and as the amount relative to *Ubiquitin* (*Ub*) message level as determined by phosphor imager readings. Probability (*P*) under the null hypothesis of no difference in message level of the maize and teosinte alleles was less than 0.01 in each case except for *Tb1/Ub* for shank (*P* < 0.05); standard error bars as shown.

ear where *tb1* is expressed, growth of the lodicules (petals) and stamens is arrested.

A previous proposal³ that *tb1* acts as a repressor of axillary meristem growth is consistent with our expression data in that the axillary organs in which *tb1* is expressed are typically reduced in size. For example, the husks of wild-type maize have reduced leaf blades and the axillary branch internodes of wild-type maize are shorter. Thus, *tb1* may act as a repressor of the growth of those organs in which the gene is expressed. *tb1* also regulates the sex of the inflorescences terminating the lateral branches and is required for the normal formation of ears. This may indicate that *tb1* has functions in addition to that of repressing organ growth.

Previously, we proposed that maize and teosinte would both carry functional alleles of *tb1* but ones whose expression is differentially regulated³. A higher level of expression for the maize allele is expected in correspondence with its shorter (more repressed) axillary branches. To test this, we quantified *tb1* mRNA levels for ear primordia from 12 plants of maize inbred W22 (carrying *Tb1 + Maize*) and 12 plants of this same inbred but with *Tb1 + teosinte* (Fig. 4). The *Tb1 + Maize* sample had a mean band intensity twice that of *Tb1 + teosinte* (*P* < 0.01), consistent with this expectation. The message level of the maize allele in immature shanks is also greater than that of the teosinte allele. The relative mRNA level for a wider sample of maize and teosinte alleles needs to be analysed to assure the generality of these results.

In view of data presented here, the following model for *tb1* in

maize evolution is proposed³. In teosinte, *tb1* is functional and is normally expressed in the secondary axillary meristems where it controls their conversion into ear shoots (including the small ear, the single husk that surrounds this ear and the single short internode that subtends it). *tb1* is not normally expressed in the primary axillary meristems of teosinte so that these are able to develop into elongated tassel-tipped branches. During the domestication of maize, humans selected an allelic variant of *tb1* that is expressed in primary axillary meristems (and probably has a high level of expression) such that these form ear shoots rather than elongated tassel-tipped branches. Thus, evolution has not proceeded by a loss/gain or change in *tb1* function, but by an alteration in gene regulation. Consistent with this view, a preliminary comparison of partial amino-acid sequences of maize and teosinte alleles revealed no fixed amino-acid differences between them, suggesting that a change in protein function has not occurred (see Supplementary Information).

The cloning of *tb1* provides a new tool not only for understanding maize evolution but also for the study of plant development. *tb1* is involved in controlling apical dominance, inflorescence development and sex determination, processes of broad interest in plant developmental biology. It may also be possible to engineer *tb1* or *tb1*-like genes to provide increased apical dominance in other crops. Using the tools of molecular genetics, it will be possible to determine whether *tb1*-like genes have played a role in the domestication of other crops, such as sunflower, that show increased apical

dominance relative to their wild ancestors¹. Comparative QTL mapping¹² suggests that this is not an unlikely prospect. □

Methods

Plant materials. The W22 line carrying *Tbl+ teosinte* was constructed as previously described³ except with six generations of backcrossing. The *tbl1-ref* stock was provided by C. Burnham and the *Mu* stocks by V. Chandler. The *Mu* tagging population, the F₂ population segregating for *tbl1-ref*, and plants for quantification of *tbl1* message levels were grown at the University of Minnesota Agricultural Experiment Station in St Paul during the summers of 1994, 1995 and 1996, respectively.

Nucleic acid analysis. DNA extraction and Southern hybridizations were done as previously described⁵. Molecular marker loci closely flanking *tbl1* were bcd1072, np1615, umc107 and umc140. Genomic DNA restriction fragments carrying the cosegregating *Mu3* element were cloned into either λ -Dash or λ -ZAP (Stratagene), following manufacturer's instructions. The cDNA library, constructed in λ -ZAP Express (Stratagene) from RNA isolated from immature ears of Pioneer line AP9, was provided by S. Briggs and T. Helentjaris (Pioneer Hi-Bred). Total cellular RNA was isolated using TRI reagent (Molecular Research Center) according to manufacturer's instructions. Northern blot analyses were performed using 10 μ g total RNA per lane as described elsewhere¹³. RNA loading was quantified by stripping the membranes and reprobing with a 0.65 kb *PstI*-*SacI* subclone of a maize *Ubiquitin* cDNA probe¹⁴. *tbl1* and *Ubiquitin* expression for the W22: *Tbl+ Maize* and W22: *Tbl+ teosinte* samples was determined by scanning northern blots with a Molecular Dynamics Storm (Phosphor) imager. Ear primordia used in the quantitative analyses (Fig. 4) were staged by length, and there was no difference between the mean lengths of the maize (29.3 mm $\pm$ 1.8) and teosinte (29.4 mm $\pm$ 1.6) genotypes.

PCR was performed using PCR Supermix (Life Technologies) with 35 cycles of 1 min at 95°C, 1 min at 58°C and 3 min at 72°C followed by 10 min at 72°C. DNA substrates were from plants carrying the three *tbl1-mum* alleles. Each reaction included the *Mu* primer (CCAACGCCAWSGCTCCATTTCGTCGA ATCC) and one of the following *tbl1*-specific primers (JD72: GGTCACTTGT AGCCAATAGC, JD76: CTTACCTGCTGATCTATTGC, JD83: GGTCAGATA GTAAGTTGTGC, JD88: TAATACATCATCATGCGAT, JD96: TCCCATCAGT AAAGCACATG). Primers JD72 and JD76 were also used to amplify a 1,500 bp portion of the *tbl1* gene from maize inbred A619. PCR products were cloned using the TA Cloning Kit version D (Invitrogen) following manufacturer's recommendations. DNA sequencing was done by the Advanced Genetics Analysis Center (University of Minnesota).

Received 19 December 1996; accepted 11 February 1997.

- Harlan, J. R. *Crops and Man* (Amer. Soc. Agron., Madison, WI, 1992).
- Ilitis, H. H. From teosinte to maize: the catastrophic sexual transmutation. *Science* **222**, 886–894 (1983).
- Doebley, J., Stec, A. & Gustus, C. *Teosinte branched1* and the origin of maize: evidence for epistasis and the evolution of dominance. *Genetics* **141**, 333–346 (1995).
- Luo, D. et al. Origin of floral asymmetry in *Antirrhinum*. *Nature* **383**, 794–799 (1996).
- Doebley, J. & Stec, A. Inheritance of the morphological differences between maize and teosinte: comparison of results for two F₂ populations. *Genetics* **134**, 559–570 (1993).
- Burnham, C. Teosinte branched. *Maize Genet. Coop. Newslett.* **33**, 74 (1959).
- Schnable, P. *tb* may condition a semi-dominant effect on tiller number. *Maize Genet. Coop. Newslett.* **66**, 5 (1992).
- Woodman, J. C. & Kremer, D. A. in *Plant Genetics* (ed. Freeling, M.) 834–836 (Liss, NY, 1985).
- Chandler, V. L. & Hardeman, K. J. The *Mu* elements of *Zea mays* Adv. *Genet.* **30**, 77–122 (1992).
- Briggs, S. & Beavis, W. in *The Maize Handbook* (eds Freeling, M. & Walbot, V.) 653–659 (Springer, New York, 1993).
- Coen, E. S. & Meyerowitz, E. M. War of the whorls: genetic interactions controlling flower development. *Nature* **353**, 31–37 (1991).
- Paterson, A. et al. Convergent domestication of cereal crops by independent mutations at corresponding genetic loci. *Science* **269**, 1714–1718 (1995).
- Sambrook, J., Fritsch, E. M. & Maniatis, T. *Molecular Cloning: a Laboratory Manual* (Cold Spring Harbor Laboratory Press, New York, 1989).
- Christiansen, A. H. & Quail, P. H. Sequence analysis and transcriptional regulation by heat shock of polyubiquitin transcripts from maize. *Plant Mol. Biol.* **12**, 619–632 (1989).

Supplementary Information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from Mary Sheehan at the London editorial office of Nature.

Acknowledgements. We thank J. Dorweiler, S. Hake and S. White for critical review of the manuscript and V. Chandler for advice on *Mu* tagging, gift of the *Mu* stocks and *Mu* element clones. This work was supported by NSF, USDA and the Agricultural Experiment Station, John Hall Fund, Graduate School and Plant Molecular Genetics Institute of the University of Minnesota. L.H. was supported by USDA funds to S. Hake.

Correspondence and requests for materials should be addressed to J.D. (jdoebley@maroon.tc.umn.edu).

A common precursor for primitive erythropoiesis and definitive haematopoiesis

Marion Kennedy*†, Meri Firpo*†‡, Kyunghye Choi†‡, Charles Wall†, Scott Robertson†, Neil Kabrun† & Gordon Keller†

* These authors contributed equally to this work.

† The National Jewish Medical and Research Center, Denver, Colorado 80206, and the Department of Immunology, University of Colorado Health Sciences Center, Denver, Colorado 80262, USA

The generation of blood cells, haematopoiesis, in the mouse embryo begins with the development of primitive nucleated erythroid cells in the yolk sac followed by the appearance of precursors for multiple definitive haematopoietic lineages^{1–4}. The later developing lineages arise from multipotential stem cells^{5,6}, but the relationship of primitive erythroid cells to these other haematopoietic populations is unknown. Using an *in vitro* embryonic stem (ES) cell differentiation system⁷, we show that primitive erythrocytes and other haematopoietic lineages arise from a common multipotential precursor that develops within embryoid bodies generated from differentiated ES cells. In response to vascular endothelial growth factor and c-kit ligand these precursors give rise to colonies containing immature cells (blasts) expressing marker genes characteristic of haematopoietic precursors. Many blast colonies also expressed β H1 and β major globins but not Brachyury, a mesodermal marker. Kinetic analysis demonstrated that the blast colony-forming cells represent a transient population, preceding the establishment of the primitive erythroid and other lineage-restricted precursors. This precursor population may represent the earliest stage of embryonic haematopoietic commitment.

Previous studies have demonstrated that the primitive erythroid lineage is established within developing embryoid bodies (EBs) by day 4 of differentiation⁸. To identify earlier precursor populations, the current series of experiments focused specifically on EBs differentiated for shorter periods of time. When cells from EBs that had been differentiated for 3 to 3.5 days were plated in methyl cellulose cultures in the presence of vascular endothelial growth factor (VEGF) and c-kit ligand (KL), colonies consisting of readily identifiable cells with an undifferentiated or blast morphology developed within 2 to 3 days (Fig. 1a, b). Secondary EBs, which were also present in these cultures, were distinguished from the blast cell colonies by the fact that they are made up of densely packed cells (Fig. 1a). Additionally, the cells within the EBs tend to have a slightly higher cytoplasm to nucleus ratio than those found in the blast colonies (Fig. 1b).

When individual blast cell colonies (4–6 days old) were picked and replated into secondary methyl cellulose cultures containing a broad spectrum of cytokines, colonies of many different haematopoietic lineages developed. The most striking pattern observed was that cultures containing colonies of primitive erythroid cells (Ery^P; Fig. 1c, d) also contained colonies of other haematopoietic lineages including definitive erythroid (Ery^D; Fig. 1e, f), multilineage (Fig. 1g, h) and myeloid (not shown). The erythroid populations were identified based on the fact that Ery^P cells are large, nucleated and express β H1 globin whereas the Ery^D cells are smaller, enucleated and express little if any embryonic globin^{9,10}. When scoring cultures, individual Ery^P and Ery^D colonies were distinguished by the size of

‡ Present addresses: DNAX, Palo Alto, California (M.F.); University of Maryland at Baltimore, Maryland, USA (K.C.).

the cells within the colonies and by the fact that Ery^p colonies were a more brilliant red colour than the Ery^d colonies. These criteria proved to be very reliable as only colonies identified as Ery^p contained cells with a primitive erythroid morphology that expressed readily detectable levels of β H1 globin (Fig. 2). Both types of colonies were found to express β -major globin. At present it is not clear if the β -major globin expression observed in the Ery^p colonies reflects expression of the adult globin in these embryonic cells or if these colonies contained small numbers of Ery^d cells.

As indicated above, the blast cell colonies grew in response to VEGF and KL. However, in cultures with low cell numbers blast colonies developed poorly, suggesting that their growth was augmented by factors produced by EB-derived subpopulations. To address this further, media conditioned by EBs as well as EB-derived cell lines were added to the blast cell cultures. Although EB-conditioned medium showed some growth enhancement of blast colonies (not shown), D4T CM, a medium conditioned by an EB-derived endothelial cell line, D4T, reproducibly showed the most significant growth potentiation, and this effect was most dramatic in low-density cultures (Fig. 3a). In contrast, D4T CM did not enhance the number of secondary EBs that developed.

In the presence of D4T CM, VEGF stimulated the growth of significant numbers of blast cell colonies but had no effect on

secondary EBs (Fig. 3b). VEGF was tested as our previous studies had shown that its receptor, flk-1, was expressed in early developing EBs (K.C. and G.K., unpublished results). KL alone stimulated few colonies and showed little enhancement in colony number when added together with VEGF. However, the combination of KL and VEGF did lead to the development of larger colonies that contained a higher proportion of Ery^d and myeloid precursors than those grown in the presence of VEGF alone (see below). Other cytokines including interleukin-6 (IL-6), IL-3 and IGF-1 stimulated few blast cell colonies on their own and showed little effect when added together with VEGF or VEGF/KL (not shown). These findings demonstrate a unique role for VEGF in the growth and development of these early embryonic haematopoietic cells.

A majority (74%) of the blast colonies generated in the presence of VEGF/KL replated to give rise to secondary haematopoietic colonies (Fig. 3c). From a total of 225 replated colonies in 9 independent experiments, 170 had haematopoietic potential. Of these, 25% gave rise to 2–10 secondary colonies, 63% gave rise to 10–50 secondary colonies and 12% yielded greater than 50 secondary colonies. Of the colonies that did replate, 33% generated primitive erythroid colonies plus colonies of other haematopoietic lineages while the remainder showed no primitive erythroid potential (Fig. 3c). Almost all (95%) secondary cultures with Ery^p

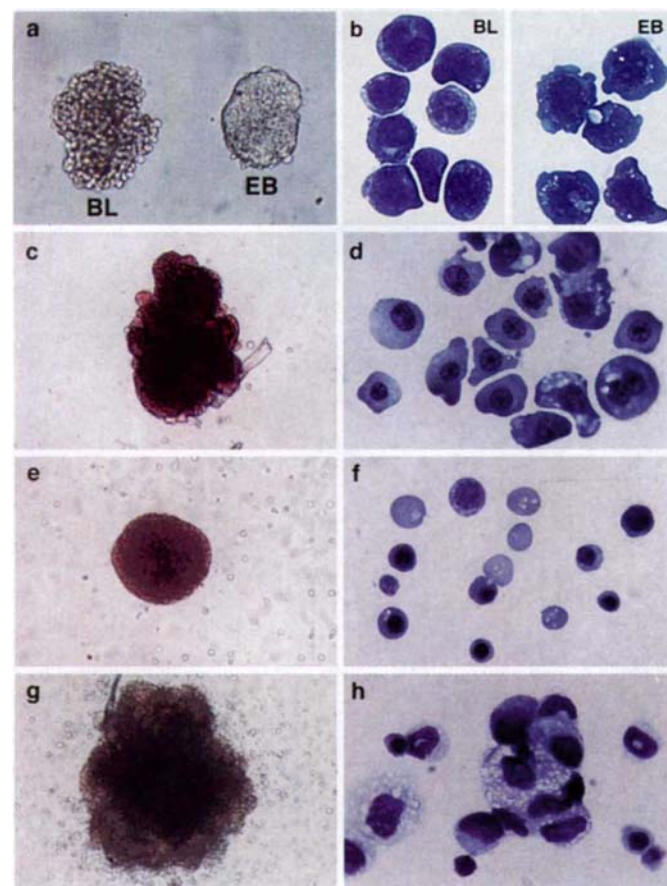


Figure 1 a–h, Photographs of colonies and corresponding cell populations. **a**, Blast cell colony (BL) and embryoid body (EB); **b**, immature cells from blast cell colonies (BL) and from secondary embryoid bodies (EB); **c**, primitive erythroid colony; **d**, primitive erythrocytes; **e**, definitive erythroid colony; **f**, definitive erythrocytes; **g**, multilineage (mix) colony; **h**, cell lineages from multilineage colony including neutrophils, macrophages and definitive erythroid cells. The primitive erythroid, definitive erythroid and multilineage colonies shown were picked from the same culture generated from a single replated blast cell colony. Cells were stained with May-Grunwald and Giemsa. Magnification for **a**, **c**, **e**: $\times 64$; for **g**: $\times 32$; and for **b**, **d**, **f** and **h**: $\times 320$.

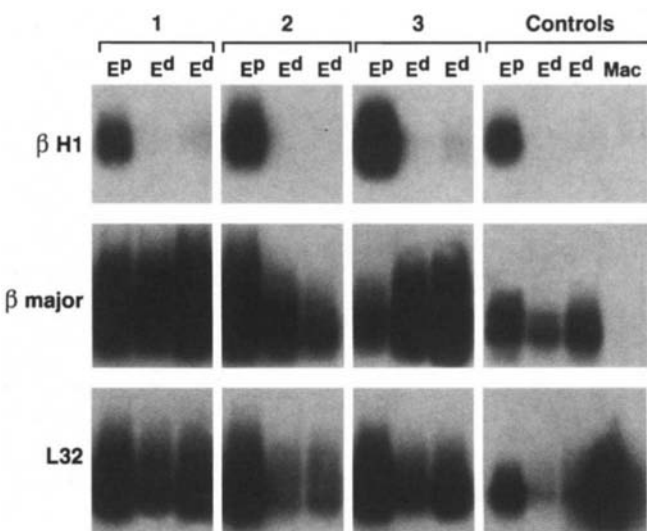


Figure 2 Globin analysis of erythroid colonies from replated cultures. The analysis of one primitive erythroid and two definitive erythroid colonies from three separate cultures is shown. Each culture was generated from a single replated blast cell colony. Control E^p represents E^p colonies grown from day 6 EB precursors, whereas the control E^d represents colonies generated from fetal liver and adult erythroid precursors, respectively. Mac represents macrophages generated from day 6 EB-derived precursors.

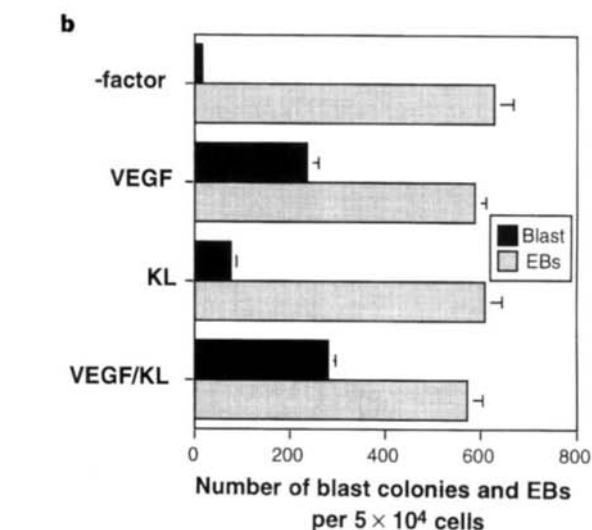
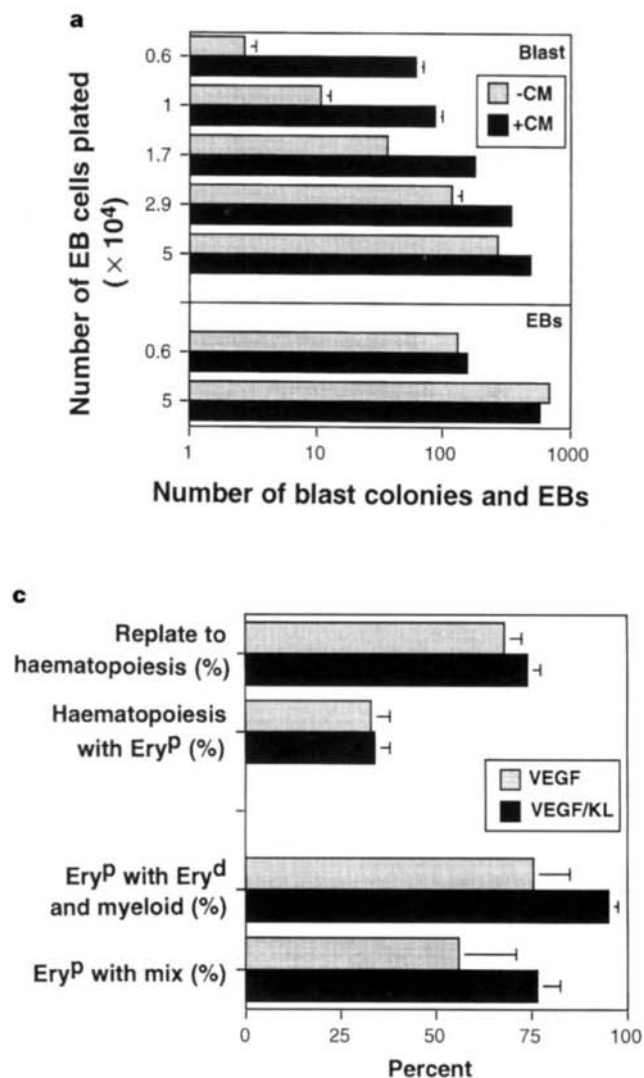


Figure 3 a, Effect of D4T-conditioned medium on blast colony and EB growth. Day 3.25 EB-derived cells were cultured at the indicated densities in the presence of VEGF, KL and Epo. Where shown, D4T CM was added at a final volume of 25%. Bars represent standard error of the mean of 3 cultures. The effect of CM was significant at all concentrations, with $P < 0.001$ in cultures with 0.6×10^4 cells and $P < 0.05$ in the culture with 5×10^4 cells. The differences in EB growth in the presence and absence of CM are not significant. CM alone had no blast colony growth promoting effect in the absence of added factors (not shown). **b**, Growth of blast cell colonies in the presence of VEGF and KL. Cells from day 3.25 EBs were cultured in the presence of VEGF, KL or both VEGF and KL and the number of blast cell colonies determined 4 days later. Bars represent standard error of the mean of 6 cultures. **c**, Analysis of blast haematopoietic potential in secondary cultures. Bars represent standard error of the mean of the per cent of the indicated correlations calculated from 5 independent experiments for VEGF-stimulated blast colonies and 9 independent experiments from blasts grown in the presence of VEGF and KL. The only significant difference in the replating potential of the blast colonies is found in the correlation of % Ery^P with Ery^d and myeloid ($P < 0.05$).

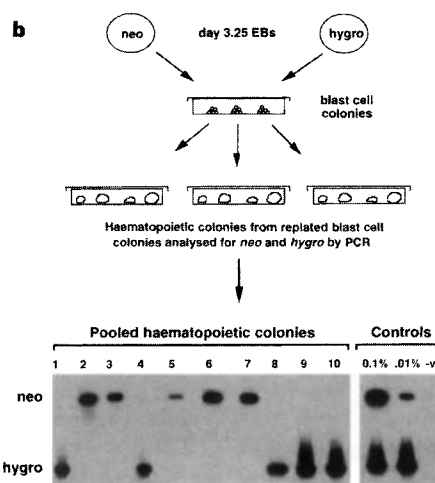
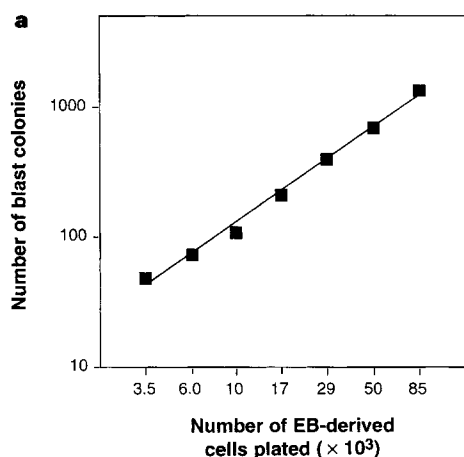


Figure 4 a, Relationship between the number of EB-derived cells plated and the number of blast cell colonies that develop. Day 3.25 EB-derived cells were cultured in the presence of VEGF/KL and 25% D4T CM. Error bars are plotted but not visible. Boxes represent the mean of four cultures at each point. **b**, PCR analysis of pools of haematopoietic colonies from blast colonies generated from

a mixture of *neo* and *hygro* containing EBs. These EBs were generated from ES cells carrying either the *neo* or *hygro* genes. Lanes 1 to 10 represent pools of haematopoietic colonies from 10 individually replated blast colonies. Controls represent 0.1% and 0.01% of the indicated population mixed with the opposite population. Minus, a sample with all the reagents but no DNA.

colonies also contained Ery^d plus myeloid colonies. A high proportion (76%) of those cultures also had multilineage (mix) colonies (Fig. 3c). Colonies grown in the presence of VEGF alone showed similar replating potential, with the exception that the number of Ery^p-containing cultures with both Ery^d and myeloid colonies was slightly lower than that observed from blast colonies grown in the presence of VEGF/KL. In contrast to the high replating potential of the blast colonies, secondary EBs showed almost no haematopoietic activity when replated under identical conditions. Out of 50 individual replated EBs, only 1 gave rise to 3 primitive erythroid colonies whereas 2 generated single myeloid/erythroid colonies.

The above replating studies demonstrate that individual blast cell colonies contain primitive erythroid precursors plus precursors of other haematopoietic lineages and support the notion that these populations share a common ancestor, provided that the colonies are clonal. As a first approach to address this issue, the relationship between the number of blast colonies and the number of EB-derived cells plated was determined. As shown in Fig. 4a, this relationship is linear with a slope of 1 suggesting that the blast colonies are of single-cell origin. In a second approach, blast cell colonies were generated from a mixture of EB-derived cells carrying genes encoding either neomycin (*neo*) or hygromycin (*hygro*) resistance (Fig. 4b). Blast cell colonies grown under these mixed culture conditions were individually replated. Pools of 10 to 15 secondary colonies derived from each primary blast colony were analysed by polymerase chain reaction (PCR) for the presence of the *neo* and *hygro* genes. The sensitivity of the PCR was designed to identify less than 0.1% of a contaminating population and as such could easily identify mixtures within this pool size. Out of 80 pools of colonies analysed from 5 different experiments, 76 were found to contain either *neo* or *hygro*, whereas 4 showed predominantly 1 marker with trace levels of the second marker. Of the 76 pools, 39 contained Ery^p colonies plus colonies of other lineages, whereas the remaining 37 had no Ery^p content. The fact that the majority of replated cultures contain only one of the markers further supports the notion that the blast cell colonies are derived from a single cell. A similar cell dose response and segregation of markers could also be observed if the blast colonies were derived from cell aggregates resulting from incomplete dissociation of the primary EB population. However,

with the high frequency of blast colonies observed, (~1 per 100 cells plated; Fig. 4a), the population would reproducibly have to contain large numbers of such aggregates. Visual inspection of the EB cells before plating indicated that the populations were virtually dissociated to single cells, with only few aggregates present in some of the experiments (<1 per 100,000 cells). Together these findings strongly suggest that the blast colonies are derived from a single cell that has the potential to generate precursors of the primitive erythroid plus other haematopoietic lineages.

Our early experiments indicated that blast cell colony-forming cells (BL-CFC) are present in day 3 to day 3.5 EBs. To define the kinetics of the development of this precursor population, EBs were collected at 6–12-h intervals between days 2.5 and 5 of differentiation and analysed for the presence of BL-CFC. The kinetics of development of the BL-CFC were compared to those of the primitive erythroid-restricted precursors (Ery^p-CFC) which were previously identified as the earliest developing haematopoietic population within the EBs⁸. As shown in Fig. 5, low numbers of BL-CFC could be detected as early as 2.5 days of differentiation, 24 h before the appearance of the first Ery^p precursors. The number of BL-CFC increased over the next 24 h, reaching maximum numbers at 3.5 days and remaining relatively constant over the next 18 h before declining sharply by day 5. Low numbers of Ery^p-CFC were detected by day 3.5 of differentiation and their number continued to increase throughout the duration of the experiment. Secondary EBs were high in the replated cultures at day 2.5 and declined steadily as the age of the primary EBs increased. These findings demonstrate that the BL-CFC represents a transient population that develops within the EBs earlier than other haematopoietic populations identified to date.

To define further the developmental status of the BL-CFC and their progeny, blast colonies were analysed for the expression of markers representing mesodermal, endothelial and early and late haematopoietic precursors. These include Brachyury¹¹, flk-1 (refs 12,13), SCL/TAL-1 (refs 14–16), CD34 (refs 17,18), GATA-1 (ref. 19), and β H1 and β -major globins. The expression pattern in the blast colonies was compared to that found in age-matched secondary EBs. As shown in Fig. 6, none of the cells within the blast colonies expressed Brachyury indicating that they do not contain mesodermal cells comparable to those found in whole EBs. In contrast, all of the blast colonies expressed flk-1, SCL/TAL-1, CD34 and GATA-1, demonstrating that they contain immature haematopoietic precursors, a finding consistent with the replating data. Many of the blast colonies also expressed β H1 and β -major globin, further supporting the notion that these colonies have the potential to generate cells of both the primitive and definitive erythroid lineages. Most of the EBs expressed Brachyury indicating that they contain mesodermal cells and as such probably represent an earlier stage of development than the blast colonies. Less than half of the EBs analysed expressed flk-1 and SCL/TAL-1, and only two showed any significant GATA-1 expression. None expressed either β H1 or β -major globins. The presence of CD34 in EBs that do not express flk-1 or SCL/TAL-1 is somewhat surprising and could indicate that expression of CD34 defines the earliest stages of haematopoietic commitment or that it is expressed in non-haematopoietic cells early in development.

This study demonstrates for the first time that the primitive erythroid and other haematopoietic lineages can arise from a common precursor, designated BL-CFC. Given their early stage of development, these precursors may have the potential to generate cells with *in vivo* haematopoietic repopulating potential, similar to those recently identified in early EBs²⁰. Our findings are not necessarily incompatible with other studies which suggest that primitive erythropoiesis and definitive haematopoiesis are generated at independent sites during embryonic development^{21–24}. It is possible that this multipotential precursor (BL-CFC) migrates to different sites in the developing embryo and generates primitive or

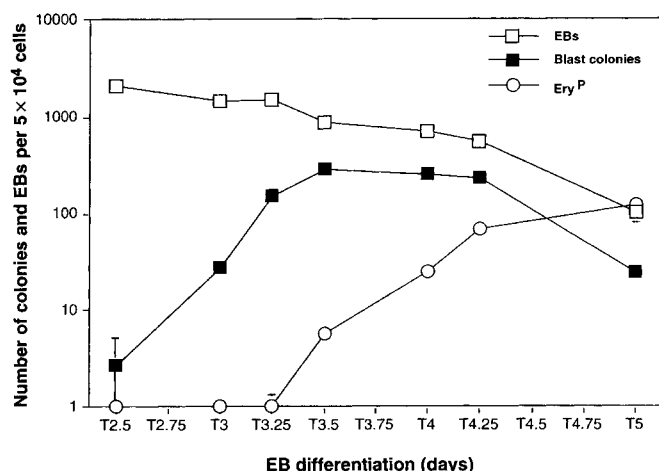


Figure 5 Kinetics of BL-CFC and Ery^p-CFC development within the EBs. EBs at the indicated stages of differentiation were collected, and the cells were dissociated and cultured in the presence of VEGF/KL for the growth of blast colonies and KL/Epo for the growth of Ery^p colonies. Bars, where visible, represent the standard error of the mean of three cultures.

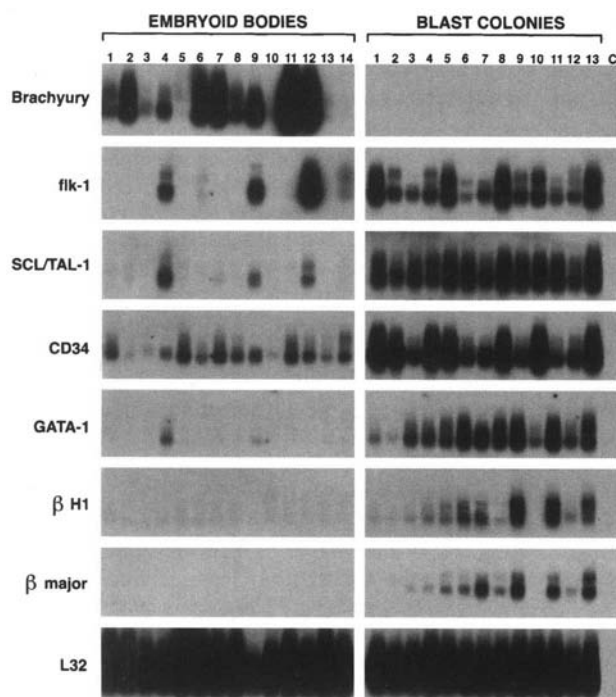


Figure 6 Expression analysis of the blast cell colonies and secondary EBs from 4-day-old replated cultures. Analysis of 13 blast colonies and 14 EBs is shown. C, Control, which consists of PCR reagents with no cells.

multilineage haematopoiesis depending upon the environment. Support for this interpretation is provided by recent studies in the *Xenopus* system which demonstrate that mesoderm of the ventral blood islands and the dorsal lateral plate of the developing embryo have the potential to generate both primitive and definitive haematopoiesis (J. B. Turpen and L. I. Zon, personal communication). Our findings differ from the recent studies of Nakano *et al.*²⁵, which indicated that the primitive and definitive erythroid lineages arise from separate precursors. These differences could reflect the fact that their analyses were carried out at later stages of ES differentiation than our studies, when the lineage-restricted precursors were already established. In addition to identifying the BL-CFC, we show that VEGF is an important regulator of this population. This finding is consistent with recent gene targeting studies demonstrating that the VEGF receptor, Flk-1, is required for the development of both the early endothelial and haematopoietic systems²⁶. The involvement of Flk-1 and VEGF in the development of the blast cell colonies suggests that the BL-CFC might also have endothelial cell potential. Indeed, our preliminary studies suggest that subpopulations of BL-CFCs are able to generate haematopoietic cells as well as cells with endothelial-like characteristics (K.C. and G.K., unpublished results). Access to these early developing precursors provided by the ES cell differentiation system will enable us to further define lineage relationships and molecular commitment steps within the embryonic haematopoietic system. □

Methods

ES cell growth and differentiation. CCE ES cells were maintained as described previously⁸. EBs were generated by plating 4.5×10^3 cells per ml in IMDM with 15% FCS, 4.5×10^{-4} M monothioglycerol (MTG), $50 \mu\text{g ml}^{-1}$ ascorbic acid and $200 \mu\text{g ml}^{-1}$ iron-saturated transferrin in 60-mm Petri dishes.

Blast colony growth. EBs collected at the indicated times following the onset of differentiation were dissociated by trypsinization and varying numbers of these cells were plated in a final volume of 1 ml methyl cellulose (1%) with 10%

FCS, 4.5×10^{-4} M MTG, $25 \mu\text{g ml}^{-1}$ ascorbic acid, $200 \mu\text{g ml}^{-1}$ iron-saturated transferrin, 25% D4T conditioned medium, VEGF (5 ng ml^{-1}) and KL (100 ng ml^{-1}) in 35-mm Petri dishes. Cultures were maintained in a humidified incubator at 37°C in an environment of 5% CO_2 in air. After 4–6 days developing blast cell colonies were picked and replated into secondary methyl cellulose cultures containing 10% plasma-derived serum (Antech, Texas), 5% protein-free hybridoma medium (PFHM2; GIBCO/BRL) and the following cytokines: KL (100 ng ml^{-1}), IL-3 (1% conditioned medium), IL-11 (25 ng ml^{-1}), GM-CSF (3 ng ml^{-1}), Epo (2 U ml^{-1}), M-CSF (5 ng ml^{-1}), G-CSF (30 ng ml^{-1}), IL-6 (5 ng ml^{-1}), LIF (1 ng ml^{-1}) and VEGF (5 ng ml^{-1}). IL-11, GM-CSF, M-CSF, IL-6, LIF and VEGF were purchased from R&D Systems. IL-3 was obtained from medium conditioned by X63 Ag8-653 myeloma cells transfected with a vector expressing IL-3 (ref. 27). Cultures were maintained as described above and secondary haematopoietic colonies were scored at 7–10 days of growth. The D4T cell line was established from day 4 EBs by infection with a retrovirus expressing the polyoma middle T gene²⁹. D4T cells express both CD31 and Flk-1, indicating that they are of the endothelial lineage. The cells are maintained in IMDM with 5% FCS, and $50 \mu\text{g ml}^{-1}$ endothelial cell growth supplement (ECGS; Collaborative Biomedical Products). Medium conditioned for 72 h on confluent cultures was collected, filtered, titred for optimal blast colony growth-stimulating potential and stored in aliquots at -80°C .

PCR analysis. Gene expression patterns in EBs and blast colonies and globin analysis in Ery^p and Ery^d colonies were determined using the global amplification strategy of Brady *et al.*²⁸. Groups of 10–50 cells from mature erythroid colonies or whole EBs or blast colonies were picked into $4 \mu\text{l}$ of first-strand buffer. Reverse transcription, tailing and PCR were performed as previously described²⁸ with the exception that the (dt)-x oligonucleotide was modified to CATCTCGAGCGGCCGC(T)₂₄. Amplified products were separated on an agarose gel and transferred to a Z-probe GT membrane (Biorad). The resulting blots were hybridized with ³²P randomly primed cDNA probes corresponding to the 3' region of the indicated genes. For analysis of βH1 expression, a probe was prepared by annealing two oligonucleotides, (5'-TGGAGTCAAAGAG-GGCATCATAGACACATGGG-3', 5'-CAGTACACTGGCAATCCCATGTG-3') which share an 8-base homology at their 3' termini. The oligo was labelled with ³²P using a Klenow fill-in reaction. All hybridizations were performed using standard methods³⁰. cDNA probes were hybridized and washed at 65°C , whereas the βH1 oligonucleotide probe was hybridized and washed at 42°C . The PCR analysis of *neo* or *hygro* in the mixing experiment was carried out on DNA prepared from replated haematopoietic colonies. The oligonucleotides used to detect *neo* were 5'-CATGATTGAACAAGATGG-3' (forward) and 5'-TCAGAAGAAGCTCGTCAAG-3' (reverse) and those for *hygro* were 5'-TGCAAGACCTGCCTGAAACC-3' (forward) and 5'-CTGCTCCATCAAGCAACCAA-3' (reverse). The expected PCR products are 760 bp for *neo* and 390 bp for *hygro*.

Received 30 December 1996; accepted 30 January 1997.

- Russell, E. S. Hereditary anemias of the mouse: a review for geneticists. *Adv. Genet.* **20**, 357–459 (1979).
- Johnson, G. R. & Barker, D. C. Erythroid progenitor cells and stimulating factors during murine embryonic and fetal development. *Exp. Hematol.* **13**, 200–208 (1985).
- Cumano, A., Curlonger, C. & Paige, C. J. Differentiation and characterization of B-cell precursors detected in the yolk sac and embryo body of embryos beginning at the 10- to 12-somite stage. *Proc. Natl Acad. Sci. USA* **90**, 6429–6433 (1993).
- Liu, C. & Auerbach, R. *In vitro* development of murine T cells from prethymic and prethymic embryonic yolk sac hematopoietic stem cells. *Development* **113**, 1315–1323 (1991).
- Keller, G. & Snodgrass, R. Life span of multipotential hematopoietic stem cells *in vivo*. *J. Exp. Med.* **171**, 1407–1418 (1990).
- Jordan, C. T. & Lemischka, I. R. Clonal and systemic analysis of long-term hematopoiesis in the mouse. *Genes Dev.* **4**, 220–232 (1990).
- Keller, G. M. *In vitro* differentiation of embryonic stem cells. *Curr. Opin. Cell Biol.* **7**, 862–869 (1995).
- Keller, G., Kennedy, M., Papayannopoulou, T. & Wiles, M. V. Hematopoietic commitment during embryonic stem cell differentiation in culture. *Mol. Cell Biol.* **13**, 473–486 (1993).
- Brotherton, T. W., Chui, D. H. K., Gaudie, J. & Patterson, M. Hemoglobin ontogeny during normal mouse fetal development. *Proc. Natl Acad. Sci. USA* **76**, 2853–2857 (1979).
- Barker, J. E. Development of the mouse hematopoietic system. I. Types of hemoglobin produced in embryonic yolk sac and liver. *Dev. Biol.* **18**, 14–29 (1968).
- Herrmann, B. G., Labeit, S., Poustka, A., King, T. R. & Lehrach, H. Cloning of the T gene required in mesoderm formation in the mouse. *Nature* **343**, 617–622 (1990).
- Yamaguchi, T. P., Dumont, D. J., Conlon, R. A., Breitman, M. L. & Rossant, J. *Flk-1*, an *flt*-related receptor tyrosine kinase is an early marker for endothelial cell precursors. *Development* **118**, 489–498 (1993).
- Millauer, B. *et al.* High affinity VEGF binding and developmental expression suggest Flk-1 as a major regulator of vasculogenesis and angiogenesis. *Cell* **72**, 835–846 (1993).
- Begley, C. G. *et al.* The gene SCL is expressed during early hematopoiesis and encodes a differentiation-related DNA-binding motif. *Proc. Natl Acad. Sci. USA* **86**, 10128–10132 (1989).

15. Robb, L. *et al.* Absence of yolk sac hematopoiesis from mice with a targeted disruption of the *scf* gene. *Proc. Natl Acad. Sci. USA* **92**, 7075–7079 (1995).
16. Shivdasani, R., Mayer, E. & Orkin, S. H. Absence of blood formation in mice lacking the T-cell leukemia oncoprotein tal-1/SCF. *Nature* **373**, 432–434 (1995).
17. Civin, C. I. *et al.* Antigenic analysis of hematopoiesis. III. A hematopoietic progenitor cell surface antigen defined by a monoclonal antibody raised against KG-1a cells. *J. Immunol.* **133**, 157–165 (1984).
18. Katz, F. E., Tindler, R., Sutherland, D. R. & Greaves, M. F. Identification of a membrane glycoprotein associated with haemopoietic progenitor cells. *Leukemia Res.* **9**, 191–198 (1985).
19. Orkin, S. GATA-binding transcription factors in hematopoietic cells. *Blood* **80**, 575–581 (1992).
20. Hole, N., Graham, G. J., Menzel, U. & Ansell, J. D. A limited temporal window for the derivation of multilineage repopulating hematopoietic progenitors during embryonic stem cell differentiation in vitro. *Blood* **88**, 1266–1267 (1996).
21. Dieterlen-Lievre, F. On the origin of haemopoietic stem cells in the avian embryo: an experimental approach. *J. Embryol. Exp. Morphol.* **33**, 607–619 (1975).
22. Kau, C. & Turpen, J. B. Dual contribution of embryonic ventral blood island and dorsal lateral plate mesoderm during ontogeny of hemopoietic cells in *Xenopus laevis*. *J. Immunol.* **131**, 2262–2266 (1983).
23. Müller, A. B., Medvinsky, A., Strouboulis, J., Grosfeld, F. & Dzierzak, E. Development of hematopoietic stem cell activity in the mouse embryo. *Immunity* **1**, 291–301 (1994).
24. Cumano, A., Dieterlen-Lievre, F. & Godin, I. Lymphoid potential, probed before circulation in mouse, is restricted to caudal intraembryonic splanchnopleura. *Cell* **86**, 907–916 (1996).
25. Nakano, T., Kodama, H. & Honjo, T. *In vitro* development of primitive and definitive erythrocytes from different precursors. *Science* **272**, 722–724 (1996).
26. Shalaby, F. *et al.* Failure of blood-island formation and vasculogenesis in Flk-1-deficient mice. *Nature* **376**, 62–66 (1995).
27. Karasuyama, H. & Melchers, F. Establishment of mouse cell lines which constitutively secrete large quantities of interleukin 2, 3, 4, or 5 using modified cDNA expression vectors. *Eur. J. Immunol.* **18**, 97–104 (1988).
28. Brady, G., Barbara, M. & Iscove, N. Representative *in vitro* cDNA amplification from individual hematopoietic cells and colonies. *Meth. Mol. Cell Biol.* **2**, 17–25 (1990).
29. Williams, R. L., Courtneidge, S. A. & Wagner, E. F. Embryonic lethality and endothelial tumors in chimeric mice expressing polyoma virus middle T oncogene. *Cell* **52**, 121–131 (1988).
30. Church, G. M. & Gilbert, W. Genomic sequencing. *Proc. Natl Acad. Sci. USA* **81**, 1991–1995 (1984).

Acknowledgements. We thank L. Carlsson, D. Ord and A. Grigoriadis for critical comments on this manuscript and L. Landskroner for assistance with graphics and photography. This work was supported by an NIH grant.

Correspondence and requests for materials should be addressed to G.M.K. (e-mail: keller@njc.org).

More hippocampal neurons in adult mice living in an enriched environment

Gerd Kempermann, H. Georg Kuhn & Fred H. Gage

The Salk Institute for Biological Studies, Laboratory of Genetics,
10010 North Torrey Pines Road, La Jolla, California 92037, USA

Neurogenesis occurs in the dentate gyrus of the hippocampus throughout the life of a rodent^{1–4}, but the function of these new neurons and the mechanisms that regulate their birth are unknown. Here we show that significantly more new neurons exist in the dentate gyrus of mice exposed to an enriched environment compared with littermates housed in standard cages. We also show, using unbiased stereology, that the enriched mice have a larger hippocampal granule cell layer and 15 per cent more granule cell neurons in the dentate gyrus.

It has long been known that experience-dependent neuroanatomical plasticity occurs in the rat brain^{5–9}. Environmental enrichment or the 'combination of complex inanimate and social stimulation'¹⁰ has been used to induce experience-dependent neuroplasticity. Morphological changes in the hippocampus include increases in hippocampal thickness^{11,12}, dendritic arborization¹³ and the number of glial cells¹². Enrichment is normally compared with standard laboratory conditions, although the type of enrichment offered in this study still represents a deprived condition compared to conditions in the wild^{10,14}. To test whether exposure to an enriched environment could lead to higher numbers of neurons in the dentate gyrus, we performed the following experiment.

Twenty-one-day-old mice were randomly distributed in two experimental groups and grew up under either standard or enriched

conditions for 40 days (Fig. 1). During the last 12 days of this period they received one daily intraperitoneal injection (50 mg per kg body weight) of bromodeoxyuridine (BrdU; Sigma). The thymidine analogue BrdU is incorporated into the DNA of dividing cells and can be detected immunohistochemically in their progeny^{4,15}. One day after the last injection, five mice from each group were perfused with 4% paraformaldehyde. The remaining animals were tested in a Morris water maze for five consecutive days; during these five days and for an additional 23 days they lived in their respective environments.

Four weeks after the last BrdU injection, the remaining mice were perfused with 4% paraformaldehyde. Immunohistochemistry for BrdU was done and the number of labelled cells was determined using unbiased counting techniques. Results revealed no statistically significant difference in the number of BrdU-positive cells in the dentate gyrus between the two groups one day after the last injection (post-injection, p.i.). This implies that enriched conditions have little or no influence on the proliferative activity of progenitor cells in the dentate gyrus. At 4 weeks p.i., however, a highly significant difference in the number of BrdU-positive cells was found (Figs 2a and 3a, b). Enriched mice had 57% more labelled cells per dentate gyrus than controls. This suggests that environmental enrichment has a survival-promoting effect on proliferating neuronal precursor cells in the dentate gyrus.

Triple labelling for BrdU, calbindin-D_{28k} (a granule cell marker¹⁶) and glial fibrillary acidic protein (GFAP, an astrocytic marker) was performed. We used confocal microscopy to determine the phenotype of the BrdU-positive cells at 4 weeks p.i. (Fig. 3c). The distribution of phenotype did not differ between the groups: 57 and 61% in controls and in enriched animals, respectively, of the BrdU-positive cells were also calbindin-positive, indicating a neuronal differentiation to a granule-cell phenotype. Multiplying the number of BrdU-labelled cells and the respective neuronal ratio leads to the conclusion that on average at least 2,490 of the new neurons per dentate gyrus that were born over 12 days (the period of BrdU injections) survived in enriched animals compared with ~1,330 new neurons in controls.

In the dentate gyrus of control animals, 16% of the BrdU-positive cells were also GFAP-positive, whereas 27% were neither calbindin- nor GFAP-positive. The respective numbers for the enriched group are 11 and 29%. Thus enriched mice showed on average ~450 surviving new astrocytes in the dentate gyrus versus ~380 in controls, consistent with earlier reports that enriched housing causes an increase in glial cells^{6,12}.

Early reports have suggested that enriched housing produces an increased 'hippocampal depth'¹¹. We determined stereologically the absolute volume of the granule cell layer and found a statistically significant increase of ~15% in the enriched group (Fig. 2b). This change of hippocampal morphology has been attributed to a wider arborization of dendrites¹³, varying sizes of neuronal nuclei¹⁷ or more glial cells¹². A higher number of neurons as a result of enriched living could contribute to the volumetric change. When we determined stereologically the absolute number of granule cells in the dentate gyrus, we found that enriched animals contained on average more than 310,000 granule cells, as opposed to 270,000 in controls (Fig. 2c). This is an increase of more than 40,000 neurons, or at least



Figure 1 **a**, Special cage for the enriched housing of 12 mice. **b**, For comparison, a standard cage for 4 control mice is shown at the same scale. Scale bar, 25 cm.

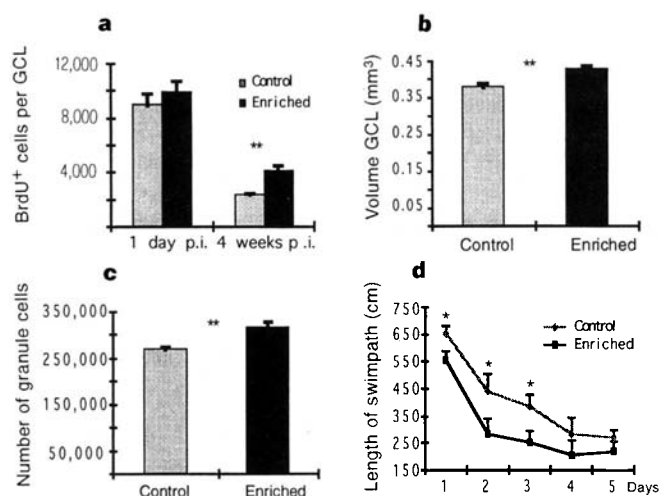


Figure 2 a–d. Four parameters in which enriched mice differ from their control littermates. **a**, Number of BrdU-positive cells per hippocampal granule cell layer (GCL). The density of BrdU-positive cells was determined stereologically (cells per mm³) and multiplied by the absolute volume of the dentate gyrus. Significance level is $P < 0.005$ (t -test). **b**, Volume of the dentate gyrus. At 4 weeks p.i., the volume of the hippocampal dentate gyrus was significantly ($P < 0.005$, t -test) increased in the enriched group. **c**, Absolute granule cell number in the dentate gyrus. At 4 weeks p.i., enriched mice had significantly more granule cell neurons ($P < 0.01$). The underlying neuronal density was 6.47 ± 0.25 per sample volume in controls and 6.56 ± 0.14 in enriched mice and thus not influenced by environmental enrichment. All data are given as mean $\pm$ standard error. N was 5 per group at 1 day p.i., and 7 per group at 4 weeks p.i. **d**, In the spatial learning task (Morris water maze), enriched mice (lower curve) have a significantly shorter swim path ($P < 0.05$, t -test) on days 1 to 3 of testing. On day 1 there was no significant difference between the groups in the first two trials, indicating equal baselines. The time to reach the platform (latency) paralleled this curve and reached the significance level on day 2. At all time points, no difference in the average swim speed could be found, and on a transfer test without a platform on day 6 no differences between groups were detected (data not shown).

15%. Although no matching stereological data for C57B16 mice are available, the baseline neuronal cell counts in the dentate gyrus falls well into the range determined for other strains¹⁸.

We conclude that environmental enrichment has a survival-promoting effect on the progeny of neuronal precursor cells in the hippocampus of mice and that these neurons add to an increased granule cell number and hippocampal volume in these animals.

The subgranular zone of the dentate gyrus is the only region of the hippocampus in which adult neurogenesis occurs. As environmental enrichment theoretically might have a stimulating effect on quiescent stem cells in other subregions, we evaluated the CA fields, but did not find BrdU-positive cells in numbers exceeding the very low and scattered ubiquitous background of dividing cells that can be found throughout the hippocampus (Fig. 3d–g). The hilar area (CA4) is an exception, as a certain rate of cell genesis, although no neurogenesis, is known to occur here^{3,4}. At 4 weeks p.i., we counted 824 ± 105 (mean $\pm$ standard error) BrdU-positive hilar cells in controls compared with 886 ± 119 in enriched animals ($P > 0.7$; t -test). The hilar volume did not change either and was 0.43 ± 0.02 mm³ in controls and 0.47 ± 0.03 mm³ in enriched animals ($P > 0.69$; t -test). These data are indicative of a rather specific effect of enriched environment on the population of newborn neurons and astrocytes in the subgranular zone.

The regulatory mechanisms underlying neurogenesis are not known and so it is unclear how an enriched environment can influence this regulation. Studies on both experience-dependent neuroplasticity⁹ and neurogenesis in the adult hippocampus¹⁹ support a role for steroid hormones in this regulation, possibly mediated through pathways involving the activation of glutamatergic receptors^{19,20}.

From *in vitro* and *in vivo* studies, it is known that trophic factors, including epidermal growth factor and fibroblast growth factor 2 (refs 21–23) can influence the fate of neuronal progenitor cells, although data on experience-dependent regulation of these factors are still lacking. It is likely that a precise temporal and spatial regulation of several factors is required for survival and differentiation of endogenous neuronal precursor cells.

To demonstrate that the environmental enrichment employed in our study was sufficient to induce the behavioural improvements in

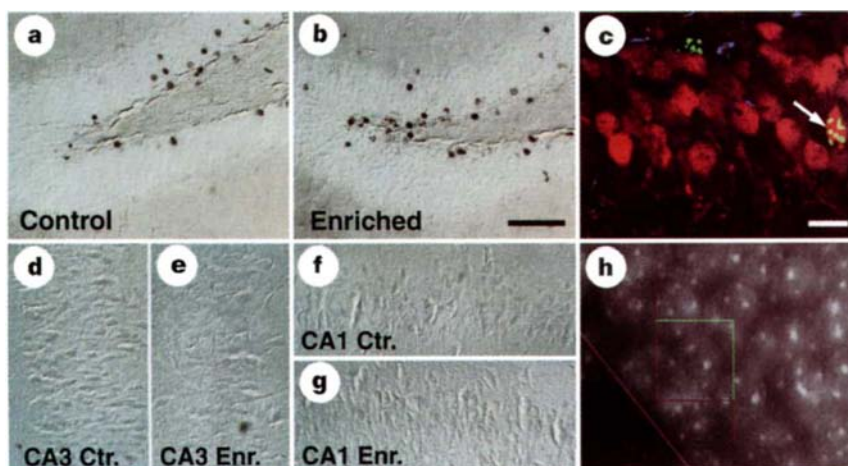


Figure 3 BrdU immunohistochemistry and phenotyping. **a** and **b**, Comparison of BrdU labelling in controls (Ctr.) and enriched (Enr.) mice at 4 weeks p.i. (for quantification, see Fig. 2a). Scale bar, 80 µm. **c**, Confocal microscopic image; arrow indicates double labelling of a BrdU-positive cell (yellow-green) with calbindin D_{28k} (red). GFAP is blue. The hilar region is at the top of the image. On the border between hilus and granule cell layer is a BrdU-positive, calbindin-negative

cell. Scale bar, 12 µm. **d–g**, Representative areas in the CA3 (**d**, **e**) and CA1 (**f**, **g**) fields mice from both groups with very few BrdU-positive cells (Scale bar in **b**, is 100 µm for **d–g**). **h**, Hoechst 33342-stained granule cells and superimposed 15 µm $\times$ 15 µm² counting frame for the stereological determination of the absolute granule cell number (see Methods).

a spatial learning task reported previously^{24,25}, we tested our animals in a Morris water maze and found a moderate, though significant improvement in mice exposed to the enriched environment (Fig. 2d). However, we cannot conclude that a greater number of neurons in the dentate gyrus leads to enhanced behavioural performance. But it is likely that a combination of increased neurons, synapses and dendrites, as well as factors still unknown, contribute to the enhanced performance induced by exposure to an enriched environment. □

Methods

Housing conditions. 24 female C57B16 mice (Harlan Sprague Dawley) were obtained at the age of weaning (21 days) and randomly distributed into two groups. Twelve mice were placed in standard conditions (4 per cage); another 12 were put in a specially designed cage of 1 m² ground area. This cage was equipped with paper tubes, nesting material, a rearrangeable set of plastic tubes, a tunnel made from sisal with various openings and a running wheel. All animals received food and water *ad libitum*. In addition to the standard food, however, animals in the 'enriched' cage received extra treats, including cheese, crackers, apples, popcorn and whole-grain nibble bars.

Behavioural testing. Mice in the 2-month survival group (4 weeks p.i.) were tested in a Morris water maze for 5 consecutive days immediately following the final BrdU injection. Mice were tested in a water maze with a computerized tracking and analysis system (San Diego Instruments) which measured the time to reach the platform (latency) and the length of the swim path. The escape platform was hidden 1 cm below the surface of the water, which had been made opaque by adding non-toxic white paint, and kept at a constant position. Each animal was tested for 4 trials per day, each lasting 40 s and beginning from 4 different start points that were randomly varied each day. If an animal did not find the platform it was set on it at the end of the trial. Animals were allowed to rest on the platform for 15 s.

Immunohistochemistry. Immunohistochemistry for the detection of BrdU requires a pretreatment of tissue sections to denature DNA¹. All staining was done on free-floating 40-µm sections. A monoclonal mouse-anti-BrdU antibody (Boehringer Mannheim; 1:400) was used in combination with avidin-biotin complex (Vector Laboratories) and a horse-anti-mouse-IgG-antibody conjugated with biotin (Vector; 1:167).

Immunofluorescent triple-labelling was done as previously described¹. Calbindin was detected by a polyclonal rabbit-anti-calbindin antibody (Swant; 1:1,000) and a Texas-red-conjugated secondary antibody (Jackson; 1:167). GFAP was detected by a polyclonal guinea-pig anti-GFAP antibody (Advanced Immuno; 1:250) and a CY-5-conjugated secondary antibody (Jackson; 1:167). All primary and secondary antibodies were applied in cock-tails with 3% horse or donkey serum and 0.1% Triton X-100 in Tris-buffered saline. To determine the phenotypes of BrdU-positive cells, 25 cells in each of 2 randomly chosen sections out of 4 were examined for each animal on a confocal microscope (Biorad/Zeiss).

Stereology. The total number of BrdU-positive cells in the subgranular layer and the corresponding sample volumes were determined in 7–9 coronal sections, 240 µm apart, containing dentate gyrus. For numbers from the hilus (CA4), we proceeded accordingly and defined as hilus that volume that is enclosed by the dentate gyrus and a virtual straight line that closes its two blades. For cell counts, the optical disector method was used to avoid oversampling errors (for reviews, see 26–28). The absolute volumes of the dentate gyrus and the hilus (CA4) were determined in a parallel haematoxylin-and-eosin-stained series of sections. For all volumetric measurements, the Cavalieri estimation was used²⁶. The absolute granule cell number was estimated using a semiautomatic stereology system, StereoInvestigator 1.0 (MicroBrightfield). For this estimation, a series of sections, 480 µm apart, was stained with Hoechst Dye 33342 (Sigma; 0.5 mg ml⁻¹ tris-buffered saline for 15 min). A 100-µm grid was projected over each section and granule cells in fields within the granule cell layer were counted in a 15 µm × 15 µm × 40 µm sample volume, disregarding the cells that were in sharp focus in the uppermost focal plane. A 60 × S-Plan-Apo oil objective with a numerical aperture of 1.40 was used on an Olympus BH-2 microscope equipped with a Panasonic CCTV video camera, and all counts were done on the video image (Fig. 3h). The resulting density was multiplied by the total volume to yield the absolute cell number.

Received 12 November 1996; accepted 6 February 1997.

- Altman, J. & Das, G. D. Autoradiographic and histologic evidence of postnatal neurogenesis in rats. *J. Comp. Neurol.* **124**, 319–335 (1965).
- Kaplan, M. S. & Hinds, J. W. Neurogenesis in the adult rat: electron microscopic analysis of light radioautographs. *Science* **197**, 1092–1094 (1977).
- Cameron, H. A. *et al.* Differentiation of newly born neurons and glia in the dentate gyrus of the adult rat. *Neuroscience* **56**, 337–344 (1993).
- Kuhn, H. G. *et al.* Neurogenesis in the dentate gyrus of the adult rat: age-related decrease of neuronal progenitor proliferation. *J. Neurosci.* **16**, 2027–2033 (1996).
- Rosenzweig, M. R. *et al.* Effects of environmental complexity and training on brain chemistry and anatomy. *J. Comp. Physiol. Psychol.* **55**, 429–437 (1962).
- Altman, J. & Das, G. D. Autoradiographic examination of the effects of enriched environment on the rate of glial multiplication in the adult rat brain. *Nature* **204**, 1161–1163 (1964).
- Cummins, R. A. *et al.* Environmentally induced changes in the brains of elderly rats. *Nature* **243**, 516–518 (1973).
- Greenough, W. T. Experiential modification of the developing brain. *Am. Sci.* **63**, 37–46 (1975).
- Meaney, M. J. *et al.* Effect of neonatal handling on age-related impairments associated with the hippocampus. *Science* **239**, 766–768 (1988).
- Rosenzweig, M. R. *et al.* Social grouping cannot account for cerebral effects of enriched environments. *Brain Res.* **153**, 563–576 (1978).
- Rosenzweig, M. R. Environmental complexity, cerebral change, and behavior. *Am. Psychol.* **21**, 321–332 (1966).
- Walsh, R. N. *et al.* The effects of environmental complexity on the histology of the rat hippocampus. *J. Comp. Neurol.* **137**, 361–366 (1969).
- Fiala, B. A. *et al.* Environmental complexity modulates growth of granule cell dendrites in developing but not adult hippocampus of rats. *Exp. Neurol.* **59**, 372–383 (1978).
- Cummins, R. A. *et al.* A developmental theory of environmental enrichment. *Science* **197**, 692–694 (1977).
- del Rio, J. A. & Soriano, E. Immunocytochemical detection of 5'-bromodeoxyuridine incorporation in the central nervous system of the mouse. *Dev. Brain Res.* **49**, 311–317 (1989).
- Sloviter, R. S. Calcium-binding protein (calbindin-D28k) and parvalbumin immunocytochemistry: localization in the rat hippocampus with specific reference to the selective vulnerability of hippocampal neurons to seizure activity. *J. Comp. Neurol.* **280**, 183–196 (1989).
- Walsh, R. N. & Cummins, R. A. Changes in hippocampal neuronal nuclei in response to environmental stimulation. *Int. J. Neurosci.* **9**, 209–212 (1979).
- Wimer, R. E. *et al.* The genetic organization of neuron number in the granule cell layer of the area dentata in house mice. *Brain Res.* **157**, 105–122 (1978).
- Cameron, H. A. *et al.* Regulation of adult neurogenesis by excitatory input and NMDA receptor activation in the dentate gyrus. *J. Neurosci.* **15**, 4687–4692 (1995).
- McEwen, B. S. Gonadal and adrenal steroids regulate neurochemical and structural plasticity of the hippocampus via cellular mechanisms involving NMDA receptors. *Cell. Mol. Neurobiol.* **16**, 103–116 (1996).
- Reynolds, B. A. *et al.* A multipotent EGF-responsive striatal embryonic progenitor cell produces neurons and astrocytes. *J. Neurosci.* **12**, 4565–4574 (1992).
- Palmer, T. D. *et al.* FGF-2-responsive neuronal progenitors reside in proliferative and quiescent regions of the adult rodent brain. *Mol. Cell. Neurosci.* **6**, 474–486 (1995).
- Craig, C. G. *et al.* In vivo growth factor expansion of endogenous subependymal neural precursor cell populations in the adult mouse brain. *J. Neurosci.* **16**, 2649–2658 (1996).
- Wainwright, P. E. *et al.* The effects of dietary fatty acid composition combined with environmental enrichment on brain and behavior in mice. *Behav. Brain Res.* **60**, 125–136 (1994).
- Pacteau, C. *et al.* Early rearing environment and dorsal hippocampal ibotenic acid lesions: long-term influences on spatial learning and alternation in the rat. *Behav. Brain Res.* **34**, 79–96 (1989).
- Gundersen, H. J. *et al.* The new stereological tools: disector, fractionator, nucleator and point sampled intercepts and their use in pathological research and diagnosis. *Acta Pathol. Microbiol. Immunol. Scand.* **96**, 857–881 (1988).
- West, M. J. New stereological methods for counting neurons. *Neurobiol. Aging* **14**, 275–285 (1993).
- Coggeshall, R. E. & Lekan, H. A. Methods for determining numbers of cells and synapses: a case for more uniform standards for review. *J. Comp. Neurol.* **364**, 6–15 (1996).

Acknowledgements. We thank Th. Sylvinia and A. Smith and their coworkers in the Animal Research Facility of the Salk Institute for their support of this study. D. A. Peterson for help with stereological questions, and M. L. Gage, I. Fisher and Th. Palmer for comments on the manuscript. G.K. was supported by the Deutsche Forschungsgemeinschaft and H.G.K. by the Hereditary Disease Foundation. This work was funded by grants from NINDS, NIA, APA and ISRT.

Correspondence and requests for materials should be addressed to F.H.G. (e-mail: fgage@salk.edu).

Deactivation and reactivation of somatosensory cortex after dorsal spinal cord injury

Neeraj Jain, Kenneth C. Catania & Jon H. Kaas

Department of Psychology, 301 Wilson Hall, Vanderbilt University, Nashville, Tennessee 37240, USA

Sensory stimuli to the body are conveyed by the spinal cord to the primary somatosensory cortex. It has long been thought that dorsal column afferents of the spinal cord represent the main pathway for these signals^{1–3}, but the physiological and behavioural consequences of cutting the dorsal column have been reported to range from mild and transitory^{4–8} to marked^{9–13}. We have re-examined this issue by sectioning the

dorsal columns in the cervical region and recording the responses to hand stimulation in the contralateral primary somatosensory cortex (area 3b). Following a complete section of the dorsal columns, neurons in area 3b become immediately and perhaps permanently unresponsive to hand stimulation. Following a partial section, the remaining dorsal column afferents continue to activate neurons within their normal cortical target territories, but after five or more weeks the area of activation is greatly expanded. After prolonged recovery periods of six months or more, the deprived hand territory becomes responsive to inputs from the face (which are unaffected by spinal cord section). Thus, area 3b of somatosensory cortex is highly dependent on dorsal spinal column inputs, and other spinal pathways do not substitute for the dorsal columns even after injury.

The full impact of dorsal column lesions on sensory processing^{4,5,9,12} and behaviour^{6,8,10,11,13} has been difficult to evaluate

for several reasons. Most important, lesions often were or may have been incomplete, and the roles of preserved inputs were not determined (reviewed in refs 1–3, 7). Spared inputs would probably gain cortical territory over time^{14,15}, thereby amplifying their significance. In addition, recording sites were often not precisely localized to specific cortical areas or parts of the somatotopic representations within those fields. Dorsal column section may affect the separate fields of somatosensory cortex differently, and somatotopic reorganization of fields would not be detected without accurate reference to the normal representational maps within those fields.

Our goal was to investigate the effects of high cervical section of the dorsal columns in monkeys on the responsiveness of neurons in the lateral part of area 3b that normally represents inputs from the hand, while considering the role of preserved inputs. Area 3b is the primary cortical target of cutaneous afferents, and is the probable

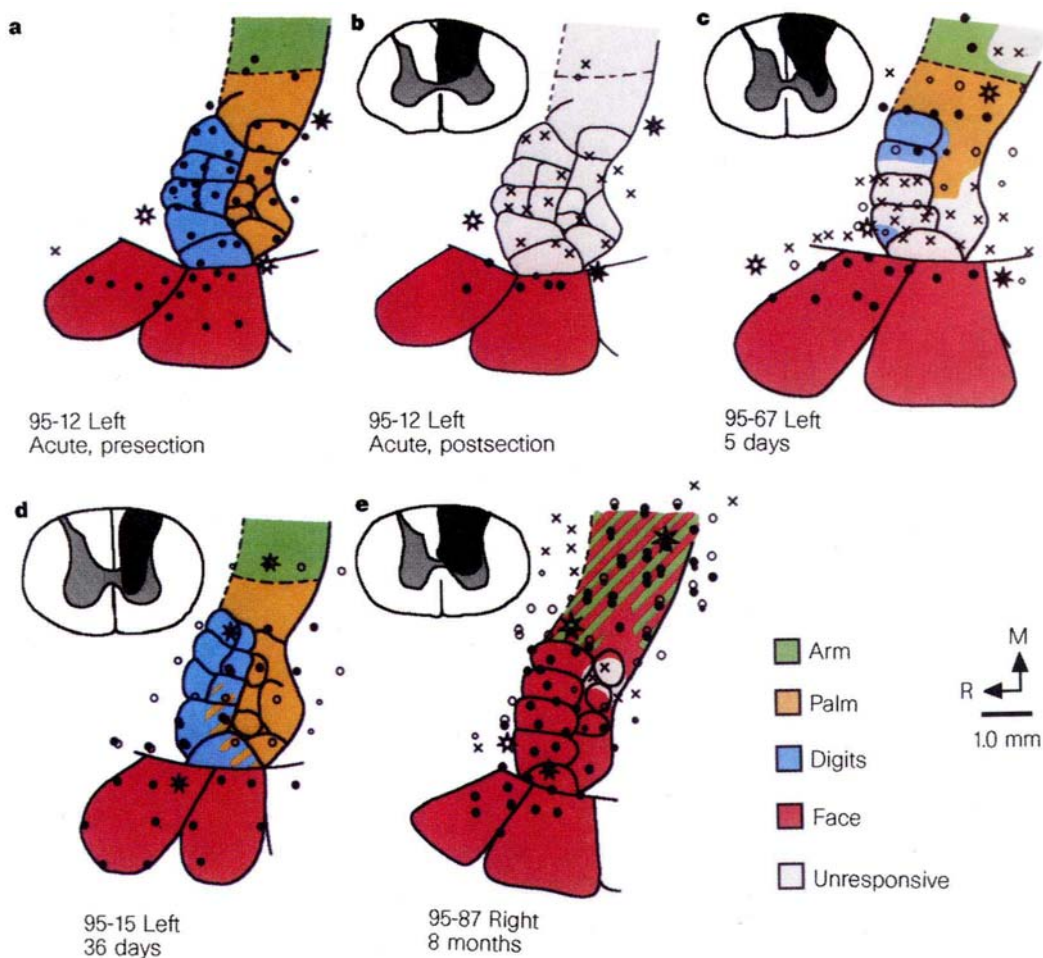


Figure 1 Effects of unilateral dorsal column section on the hand, face and arm regions of contralateral primary (area 3b) somatosensory cortex. **a**, Normal map of an adult owl monkey in area 3b before, and **b**, immediately after a complete contralateral dorsal column section. No responses were evoked in the region of normal hand representation after the section. **c**, Five days after a partial dorsal column section, the deafferented portions of the hand representation were still unresponsive. The medial region with normal responses corresponds to remaining dorsal column fibres (see text and Fig. 2). **d**, Thirty-six days after partial section there is an expansion of the cortical territory of the preserved dorsal column inputs from the deafferented hand. Digits 4 and 5 and adjoining palm activate all of the hand cortex. In the medial region, corresponding to the preserved inputs, the neural responses were normal, whereas in the lateral regions the receptive fields were large and responses highly abnormal. **e**, Eight

months after complete section, hand and arm cortex respond to stimulation of chin. Cortex responsive to the arm (green) overlaid that responsive to the chin and also showed some enhancement in cortical territory. In each case, the boundaries of area 3b (outlined coloured regions) were reconstructed and subdivisions related to upper and lower face (red), the five digits (blue), and palm (yellow) were identified in myelin-stained sections cut parallel to the pial surface¹⁷. The border of the arm representation (green) was estimated from recordings in normal animals. Recording sites are marked by black dots (neurons responding to light touch or hair displacement), unfilled dots (neurons responding to taps), or crosses (neurons were unresponsive). Stars indicate the location of marker lesions. Insets show reconstructions of the extents of the dorsal column lesions (black) for each case. M, Medial; R, rostral.

homologue of S1 in non-primate mammals¹⁶. Lesions at the C3/C4 level are high enough to disrupt afferents from all of the hand (and lower body) while leaving intact some inputs from the arm as well as those from face. We were able to relate recording sites to precise locations in area 3b using marker microlesions, later viewed in sections of flattened cortex cut parallel to the pial surface and processed for myelin. In such brain sections, area 3b stands out as a densely myelinated strip of cortex with sharp borders. Furthermore, we made the fortuitous discovery that narrow myelin-light septa divide the myelin-dense hand representation in area 3b into sub-regions each activated by a glabrous digit or palmar pad of the hand¹⁷. Thus, we could relate recording sites to precise locations within the normal map of the hand to see if neurons were activated from the original or different parts of the body surface. Finally, we included an additional and highly effective procedure to evaluate the effectiveness of dorsal column lesions. After the section, but before recording, we injected the transganglionic tracer, B-HRP, at multiple (18–20) sites into the skin of the hand. This tracer does not interfere with the responsiveness of peripheral afferents, but is transported by intact afferents that terminate onto second-order neurons in the cuneate nucleus of the brain stem^{18,19}. Any trace of transported B-HRP in the cuneate nucleus was taken as evidence that some afferents remained intact. Moreover, because these afferents terminate in precise somatotopic order in the pars rotunda of the cuneate nucleus²⁰, we could determine the source of any remaining afferents in the dorsal columns.

There are three important conclusions from our experiments. The first is that complete high cervical dorsal column section produces a complete deactivation of hand cortex in area 3b. Part of the evidence for this came from two monkeys (95-12 and 95-17) in which standard microelectrode multiunit recordings were made from the hand and adjoining face and arm portions of area 3b before and immediately after dorsal column section. Before section, neurons were highly responsive to light punctate contact on the skin or hair displacement, as previously shown¹². Receptive field locations reflected the normal somatotopic organization of area 3b (Fig. 1a, and case 95-17 (data not shown)), and matched the isomorph of the hand subsequently revealed by the myelin pattern in brain sections. Immediately after the contralateral dorsal columns section, in both cases neurons throughout the hand representation were completely unresponsive to any type of cutaneous or deep receptor stimulation on the hand or other parts of the body (Fig. 1b). Inputs from the face were unaffected by the dorsal column section, and they activated neurons in the face cortex as before. The reconstructed lesions were judged to be complete, and they did not invade the dorsolateral spinal cord (Fig. 1b, inset).

Results obtained from a third monkey (96-53) 20 days after complete section of the dorsal columns indicate that neurons do not acquire new sources of activation within this period. As in the case of the monkeys studied immediately after lesion, no neurons were found in the hand portion of area 3b that were activated by somatosensory stimuli. The reconstructed lesion was judged to be complete, and none of the tracer injected in the hand was transported to the cuneate nucleus (see Fig. 2).

Incomplete dorsal column lesions allowed portions of hand cortex to be activated by light contact on the hand, whereas other portions of hand cortex were unresponsive. In monkey 95-67, studied five days after contralateral dorsal column section, there was a large region of unresponsive cortex in the lateral part of the hand cortex, but normal responses and receptive fields were found for neurons in parts of the morphological map related to digits 4 and 5 and the ulnar palm, as well as at one recording site for digit 2 (Fig. 1c). Reconstruction of the lesion in this monkey revealed that a small ventral portion of the dorsal column was intact (Fig. 1c, inset). More significantly, tracer (B-HRP) injected into the hand after the lesion reached medial parts of the cuneate nucleus (Fig. 2d) that are normally innervated by inputs from digits 4 and 5 and the adjacent palm²². Because these inputs activated only the appropriate parts of the hand map, there was no clear evidence for short-term plasticity. Nevertheless, these sparse remaining inputs may have come to activate more cortical neurons than normal within somatotopically appropriate parts of the cortex. In summary, recordings from these four monkeys demonstrate the essential role of the dorsal column pathway.

Our second conclusion is that preserved inputs from the hand come to activate larger than normal and somatotopically inappropriate parts of the hand representation over longer recovery periods. In case 95-15, studied only 36 days after an extensive but incomplete dorsal column lesion, the hand cortex of contralateral area 3b was completely reactivated (Fig. 1d). While neurons at some recording sites had normally located receptive fields and response properties, neurons throughout much of the hand area had large mislocated receptive fields, typically involving several digits or digits and palm. These neurons gave abnormal responses which habituated to repeated stimulation and were less easily evoked by punctate stimuli. Similar results were obtained in a second case (95-32) studied eight weeks after incomplete dorsal column section (data not shown). We can conclude that after recovery periods of five or more weeks, preserved inputs retain normal access to some cortical neurons while gaining activating access to other cortical neurons in previously unresponsive parts of the hand representation.

Our third conclusion is that a much more massive type of

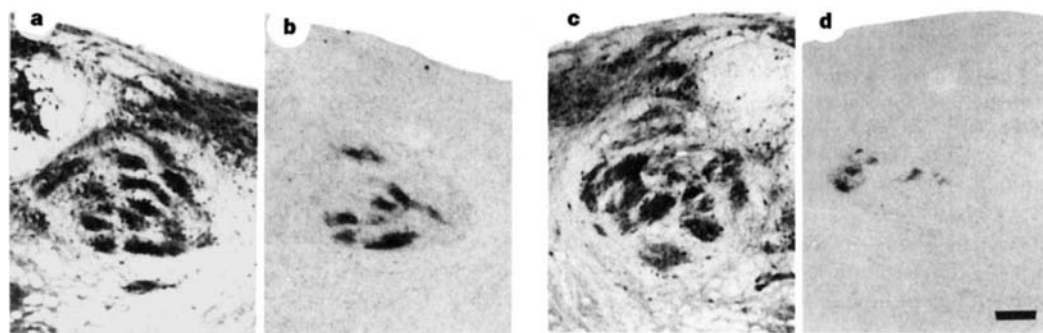


Figure 2 Transport of cholera toxin B subunit linked to horseradish peroxidase (B-HRP) injected at multiple sites in the hand to the pars rotunda of the cuneate nucleus in **b**, a normal owl monkey, and **d**, an owl monkey, 95-67, with partial dorsal column section; **a** and **c** are respective adjacent cytochrome oxidase stained sections. In normal owl monkey all of the pars rotunda is filled with label

throughout its rostrocaudal extent, whereas in owl monkey 95-67 only two sections were seen with label in the medial and mediodorsal regions after similar injections, indicating that spared fibres corresponded to few inputs from ulnar digits and palm. Dorsal is to the top. Medial is to the left for **a** and **b**, and to the right for **c** and **d**. Scale bar, 200 μ m.

reorganization can occur over even longer recovery periods of six months or more. In a monkey (95-87) studied eight months after a complete dorsal column section, no responses to stimulation of any part of the hand were found for neurons throughout the morphological map of the hand (Fig. 1e). Thus, the hand region in area 3b was not reactivated by afferents from the hand that course outside the dorsal columns. Instead, neurons throughout the hand cortex, as well as in the arm cortex, responded briskly to movement of the hairs on the chin. In addition, within the arm representation and the most medial portion of the palm cortex, neurons at the same recording sites also responded to hair displacement on the arm. As inputs from parts of the arm and the face enter the spinal cord and lower brain stem above the level of the C3/C4 spinal cord lesion, somatotopically appropriate activation from these inputs would be expected. However, these inputs clearly gained access to and reactivated neurons throughout the somatotopically inappropriate hand region of the cortex. In this case, no label from the injected hand was transported to the cuneate nucleus and the lesion did not invade the dorsolateral spinal cord (Fig. 1e inset).

Results from another monkey (95-31) studied after six months of recovery were similar, except that the lesion was incomplete and a few inputs from the hand remained. As in the previous case, neurons through most of the mediolateral extent of the hand representation were abnormally activated by inputs from the face. In addition, spared inputs from digits 4 and 5 and the adjoining palm had expanded territories within the hand representation. Thus, the face regions of cortex can expand into the deactivated hand cortex, even when some dorsal columns afferents from the hand remain.

Our results from monkeys with complete high cervical lesions of the dorsal columns show that under normal recording conditions, neurons in area 3b of somatosensory cortex completely and perhaps permanently depend on dorsal column cutaneous afferents. Thus, preserved or recovered sensory abilities after complete dorsal column lesions do not depend on area 3b. However, if any dorsal column inputs from the hand remain after the lesion, these inputs maintain access to their normal target neurons in area 3b, and subsequently, over weeks of recovery, expand their cortical territories. The greatly expanded effectiveness of even a sparse population of remaining afferents could be responsible for much of the behavioural recovery.

The second type of more massive but more slowly developing reactivation of hand cortex by face inputs could involve the growth of preserved inputs from the hand or the face into denervated parts of the cuneate nucleus. Similar reactivations of hand cortex by face inputs after amputation¹⁸ or deafferentation²³ of the arm have been reported, so this type of reactivation seems to be a general consequence of an extensive loss of forearm afferents.

Our results expand on previous reports of massive reorganization in two important ways. First, the removal of the dorsal column inputs alone is sufficient to cause expansion of the face into the hand portion of area 3b. Unlike the cases in earlier studies, where all of the sensory inputs from the arm were interrupted, spinothalamic and other inputs in the lateral funiculus remained intact in the present cases, and they did not substitute for dorsal column inputs or prevent somatotopic reorganization. Second, this large-scale reorganization can take place in a much shorter time (six months) than the years of recovery allowed in previous cases. Such massive cortical reorganization would seem to offer no benefit and might be the reason for phantom sensations reported after spinal injuries in humans^{24,25}.

Methods

Surgery. Under ketamine (15 mg kg⁻¹, intramuscular) and xylazine (0.4 mg kg⁻¹, intramuscular) anaesthesia, dorsal column fibres were sectioned at the C3/C4 level on one side in 7 adult owl monkeys (*Aotus trivirgatus*). Three days before recording from cortex, cholera toxin B subunit linked to horseradish

peroxidase (B-HRP) was injected at 18–20 subcutaneous sites on both sides of the affected hand¹⁹ under isoflurane anaesthesia. All protocols were approved by the Institutional Animal Care and Use Committee and followed NIH guidelines.

Mapping. Standard multiunit mapping methods were used^{19,21} in animals anaesthetized with ketamine and xylazine. Brushes, probes and Semmes–Weinstein filaments were used to stimulate cutaneous receptors; responses to joint movements and taps were also noted. No noxious stimuli were used. Electrode penetrations were marked on a photograph of the brain surface. Microlesions were made at reference sites so that physiological results could be accurately related to cortical histology.

Histology. After perfusion, the cortex was removed, flattened between glass slides, cryoprotected and cut parallel to the pial surface at 40 µm. The sections were stained for myelin by a modified Gallyas procedure²⁶. The medulla was cut coronally and a series of sections were processed for HRP²⁷. Another series of sections was reacted for cytochrome oxidase activity²⁸. The spinal cord was cut parasagittally at 52 µm and all sections were processed for HRP, except in the case of animals mapped immediately after the lesion, where the spinal cord was cut in a coronal plane and all sections were processed for cytochrome oxidase.

Received 30 October; accepted 6 February 1997.

- Mountcastle, V. B. Central nervous mechanisms in mechanoreceptive sensibility. In *Handbook of Physiology—The Nervous System III* (eds Darian-Smith, I.) 789–878 (American Physiological Society, Bethesda, MD, 1984).
- Tommerdahl, M. et al. Effects of spinal dorsal column transection on the response of monkey anterior parietal cortex to repetitive skin stimulation. *Cerebral Cortex* **6**, 131–155 (1996).
- Beck, C. H. M. Dual dorsal columns: a review. *Can. J. Neurol. Sci.* **3**, 1–7 (1976).
- Andersson, S. A., Finger, S. & Norrrell, U. Cerebral units activated by tactile stimuli via a ventral spinal pathway in primates. *Acta Physiol. Scand.* **93**, 119–128 (1975).
- Eidelberg, E., Kreinick, C. J. & Langescheid, C. On the possible role of afferent pathways in skin sensation. *Exp. Neurol.* **47**, 419–432 (1975).
- Glendinning, D. S. & Vierck, C. J. Jr Lack of proprioceptive deficit after dorsal column lesions in monkeys. *Neurology* **43**, 363–366 (1993).
- Wall, P. D. The sensory and motor role of impulses travelling in the dorsal columns towards cerebral cortex. *Brain* **93**, 505–524 (1970).
- Vierck, C. J. Jr Alterations of spatio-tactile discrimination after lesions of primate spinal cord. *Brain Res.* **58**, 69–79 (1973).
- Dreyer, D. A., Schneider, R. J., Metz, C. B. & Whitsell, B. L. Differential contributions of spinal pathways to body representation in post central gyrus of Macaca mulatta. *J. Neurophysiol.* **37**, 119–145 (1974).
- Glendinning, D. S., Vierck, J. C. J. & Cooper, B. Y. The effect of fasciculus cuneatus lesions on finger positioning and long latency reflexes in monkeys. *Exp. Brain Res.* **93**, 104–116 (1993).
- Glendinning, D. S., Cooper, B. Y., Vierck, C. J. Jr & Leonard, C. M. Altered precision grasping in stump-tail macaques after fasciculus cuneatus lesions. *Somatosens. Motor Res.* **9**, 61–73 (1991).
- Schneider, R. J. Loss of information concerning hair displacement and other somatic stimuli in the first somatic sensory cortex of unanesthetized macaca mulatta monkeys following dorsal funicular transections. *Exp. Brain Res.* **83**, 105–114 (1990).
- Vierck, C. J. Jr, Cohen, R. H. & Cooper, B. Y. Effects of spinal lesions on temporal resolution of cutaneous sensations. *Somatosensory Res.* **3**, 45–56 (1985).
- Kaas, J. H. Plasticity of sensory and motor maps in adult mammals. *Annu. Rev. Neurosci.* **14**, 137–167 (1991).
- Merzenich, M. M. et al. Topographic reorganization of somatosensory cortical areas 3B and 1 in adult monkeys following restricted deafferentation. *Neuroscience* **8**, 33–55 (1983).
- Kaas, J. H. What if anything is S1? The organization of the “first somatosensory area” of cortex. *Physiol. Rev.* **63**, 206–231 (1983).
- Jain, N., Catania, K. C. & Kaas, J. H. An anatomical isomorph of the hand in somatosensory cortex of owl monkeys and its immutability following peripheral deafferentation. *Soc. Neuro. Sci. Abstr.* **22**, 1054 (1996).
- Florence, S. L. & Kaas, J. H. Large-scale organization at multiple levels of the somatosensory pathway follows therapeutic amputation of the hand in monkeys. *J. Neurosci.* **15**, 8083–8095 (1995).
- Jain, N., Florence, S. L. & Kaas, J. H. Limits on plasticity in somatosensory cortex of adult rats: hindlimb cortex is not reactivated after dorsal column section. *J. Neurophysiol.* **73**, 1537–1546 (1995).
- Florence, S. L., Wall, J. T. & Kaas, J. H. Somatotopic organization of inputs from the hand to the spinal gray and cuneate nucleus of monkeys with observations on the cuneate nucleus of humans. *J. Comp. Neurol.* **286**, 48–70 (1989).
- Merzenich, M. M., Kaas, J. H., Sur, M. & Lin, C. S. Double representation of the body surface within cytoarchitectonic areas 3b and 1 in ‘S1’ in the owl monkey (*Aotus trivirgatus*). *J. Comp. Neurol.* **181**, 41–73 (1978).
- Florence, S. L., Wall, J. T. & Kaas, J. H. Central projections from the skin of the hand in squirrel monkeys. *J. Comp. Neurol.* **311**, 563–578 (1991).
- Pons, T. P. et al. Massive cortical reorganization after sensory deafferentation in adult macaques. *Science* **252**, 1857–1860 (1991).
- Sweet, W. H. ‘Phantom’ sensations following intraspinal injury. *Neurochirurgia* **18**, 139–154 (1975).
- Davis, R. W. Phantom sensation, phantom pain, and stump pain. *Arch. Phys. Med. Rehabil.* **74**, 79–91 (1993).
- Gallyas, F. Silver staining of myelin by means of physical development. *Neurol. Res.* **1**, 203–209 (1979).
- Gibson, A. R., Hansma, D. I. & Robinson, F. R. A sensitive low artifact TMB procedure for the demonstration of WGA-HRP in the CNS. *Brain Res.* **298**, 235–241 (1984).
- Wong-Riley, M. T. T. Changes in the visual system of monocularly sutured or enucleated cats demonstrable with cytochrome oxidase histochemistry. *Brain Res.* **171**, 11–28 (1979).

Acknowledgements. We thank J. Schall, S. Florence and M. Pospichal for comments on the manuscript, and J. Ives and L. Trice for technical assistance.

Correspondence and requests for materials should be addressed to J.H.K. (e-mail: jainm@ctrvax.vanderbilt.edu).

A potent and selective endogenous agonist for the μ -opiate receptor

James E. Zadina, Laszlo Hackler, Lin-Jun Ge & Abba J. Kastin

Veterans Affairs Medical Center and Tulane University School of Medicine, New Orleans, Louisiana 70146, USA

Peptides have been identified in mammalian brain that are considered to be endogenous agonists for the δ (enkephalins) and κ (dynorphins) opiate receptors, but none has been found to have any preference for the μ receptor¹⁻³. Because morphine and other compounds that are clinically useful and open to abuse act primarily at the μ receptor⁴, it could be important to identify endogenous peptides specific for this site. Here we report the discovery and isolation from brain of such a peptide, endomorphin-1 (Tyr-Pro-Trp-Phe-NH₂), which has a high affinity ($K_i = 360$ pM) and selectivity (4,000- and 15,000-fold preference over the δ and κ receptors) for the μ receptor. This peptide is more effective than the μ -selective analogue DAMGO *in vitro* and it produces potent and prolonged analgesia in mice. A second peptide, endomorphin-2 (Tyr-Pro-Phe-Phe-NH₂), which differs by one amino acid, was also isolated. The new peptides have the highest specificity and affinity for the μ receptor of any endogenous substance so far described and they may be natural ligands for this receptor.

Endorphins, enkephalins and dynorphins, which all have the amino-terminal amino-acid sequence Tyr-Gly-Gly-Phe, bind with low to moderate specificity to the three opiate receptors³ (Table 1). β -Endorphin binds to the μ and δ receptors with comparable affinity, but because of their selectivity, Met- and Leu-enkephalin are considered to be the endogenous ligands for the δ receptor and dynorphins for the κ receptor. No known mammalian peptide, however, has both high affinity and selectivity for the μ receptor; for example, Tyr-W-MIF-1 (Tyr-Pro-Trp-Gly-NH₂), a peptide in brain^{5,6} with opiate-related activity⁵⁻¹¹, is highly selective for μ rather than δ and κ receptors (>200- and 300-fold), but its affinity for the μ receptor ($K_i = 70$ nM) is relatively low¹². Natural amino-acid substitutions in the first three positions of Tyr-W-MIF-1 are not well tolerated for opiate binding, presumably because of a strict requirement for the amino and phenolic groups of Tyr at position 1, an appropriate spacer (Gly at position 2 and 3 or Pro at position 2) and an aromatic group (Phe at position 3 or 4 or Trp at position 3)¹³. With a view to identifying molecules that occur naturally in the brain, peptides containing all possible natural amino-acid substitutions at position 4 of Tyr-W-MIF-1 were synthesized by polyethylene pin technology¹⁴ and screened for opiate-receptor binding. A high-affinity, selective and biologically potent sequence was discovered and then identified in brain.

Figure 1 (top) shows the relative affinity of each peptide for the μ receptor (black bars). The peptide with Phe in the fourth position (Phe-4-Tyr-W-MIF-1) had a greatly increased affinity compared with the other peptides: its affinity estimated from two peptide sets was >50-fold greater than that of the parent compound Tyr-W-MIF-1, with Gly in position 4.

The selectivity of the peptides for μ rather than δ receptors (Fig. 1, white bars) is compared with that of Tyr-W-MIF-1, which has a selectivity ratio of 240 in these assays, in agreement with our previous report¹². This parent peptide proved to be one of the most selective in the set. The Phe-4 peptide, however, was far more selective for μ sites than the other peptides tested.

A more hydrophobic amino acid after the Trp 3 residue tended to

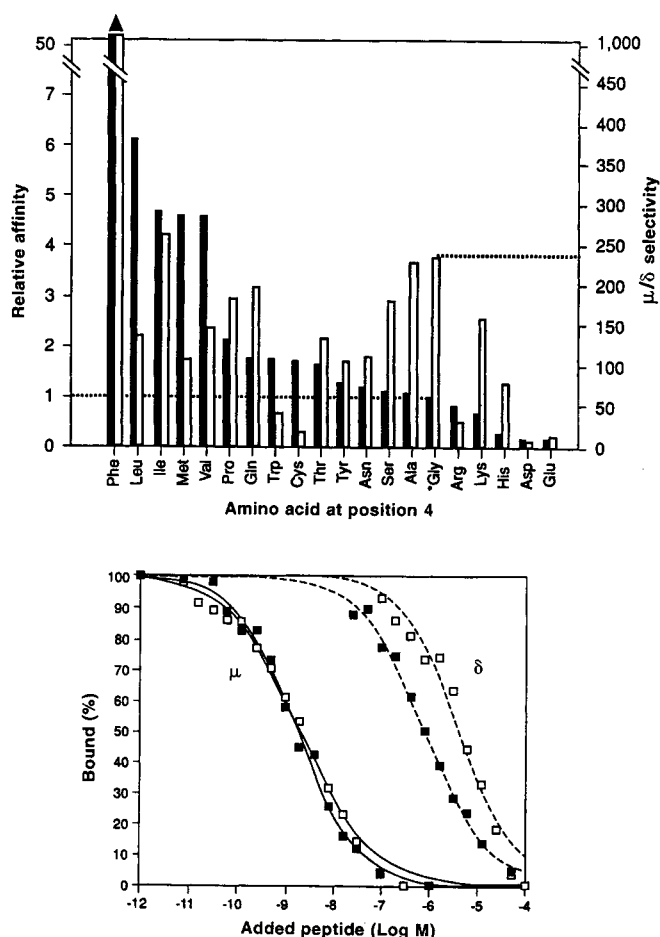


Figure 1 Binding of analogues to opiate receptors. Top, affinity and selectivity of 20 Tyr-Pro-Trp-X-NH₂ analogues. The K_i for the μ opiate receptor (black bars) is expressed relative to that of Tyr-W-MIF-1 on the left axis. The dashed line from the left axis to the bar for Tyr-W-MIF-1 (Gly at position 4; asterisk) indicates an affinity ratio of 1. The selectivity for μ over δ receptors ($\delta K_i/\mu K_i$) is shown on the right axis and indicated by white bars. The dashed line from the right axis represents the 240-fold selectivity of Tyr-W-MIF-1 for μ receptors. The arrow above the Phe analogue indicates that its extraordinary affinity and selectivity were well outside the range of the other compounds. Binding of [³H]-DAMGO (Tyr-D-Ala-Gly-N-Me-Phe-Gly-ol) to μ receptors or of [³H]-pCl-DPDPE (D-Pen², pCl-Phe⁴, D-Pen⁵-enkephalin) to δ receptors was assayed as described¹². Values represent the mean of two separate assays with each of two separate syntheses of the set of 20 peptides. Bottom, comparative binding of the Phe-4 peptide and DAMGO to μ and δ receptors. The Phe-4 peptide (white squares) inhibited μ [³H]-DAMGO (solid line) binding with an IC_{50} comparable to that of DAMGO (black squares), but it inhibited δ [³H]-pCl-DPDPE (dotted line) binding with a higher IC_{50} , reflecting its greater selectivity for the μ receptor. K_i values are given in Table 1.

increase μ receptor binding. The correlation between half-maximal inhibition (IC_{50}) of μ binding by synthetic peptide and the hydrophobicity¹⁵ of uncharged amino acids in position 4 was significant ($P < 0.001$). Based on the r^2 value (0.56), however, only about half of the variance in binding affinity was attributable to hydrophobicity, and the binding of the Phe-4 analogue was >2 standard error units stronger than predicted from the regression. This indicates that physical chemical factors in addition to hydrophobicity are important for its high binding affinity.

Figure 1 (bottom) and Table 1 show that, in comparison with one of the most potent and μ -selective analogues of enkephalin available (DAMGO), the Phe-4 peptide has an equal affinity and greater

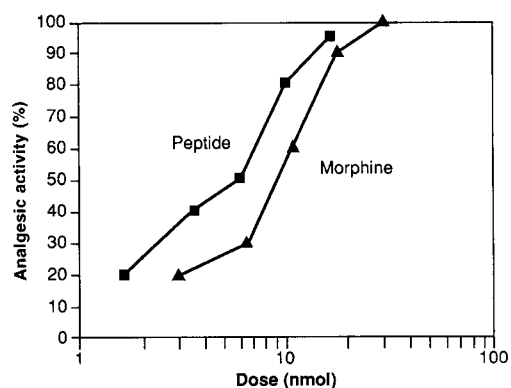
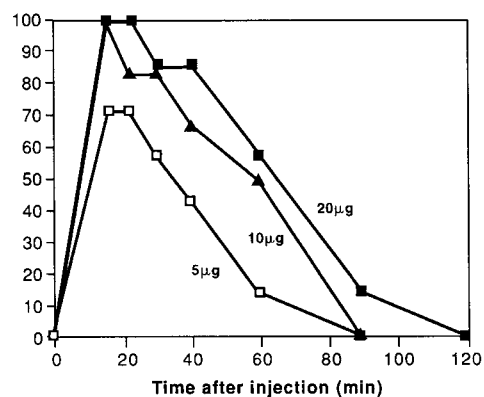


Figure 2 Left, dose-response curves for the analgesic effects of morphine and the Phe-4 peptide 15 min after i.c.v. administration to mice ($n = 10$), where dosage is in nmol per animal. Analgesia was defined as a doubling from baseline of the tail-flick latency in response to a focused light. Data were analysed quantally^{2,30} with the BLISS program. The half-maximal effective dose (ED_{50}) and 95%



confidence intervals for the peptide were 4.7 nmol (3.1–6.7) and for morphine 7.5 nmol (4.8–10.5). Right, time course of the analgesic effects of the Phe-4 peptide. Mice were given 5, 10 or 20 µg Phe-4 peptide i.c.v. and tested at various times after injection.

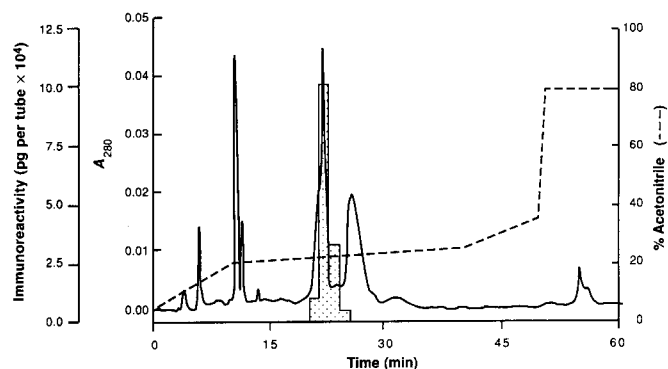


Figure 3 Isolation of Tyr-Pro-Trp-Phe-NH₂ from bovine frontal cortex. The immunoreactive (dotted bar) and ultraviolet absorbance profiles are shown from the final reversed-phase HPLC purification step. The peak of immunoreactive material isolated at fraction 24 was subjected to Edman degradation for sequencing.

selectivity for the μ receptor. Table 1 also shows that the Phe-4 peptide has higher affinity and selectivity for the μ receptor than the three principal endogenous mammalian peptides with opiate activity.

The analogue with Phe at position 4 had potent μ -selective activity *in vitro*, inhibiting electrically induced contractions of the guinea-pig ileum¹⁶ with an IC_{50} significantly more potent than that of DAMGO (3.6 ± 0.3 nM compared with 6.8 ± 0.8 nM, mean $\pm$ s.e.m. of four assays; $F_{1,6} = 13.4$, $P < 0.05$). The effect of the Phe-4 peptide was blocked and reversed by the antagonists naloxone (0.5 μ M) and the μ -selective CTOP (D-Phe-Cys-Tyr-D-Trp-Orn-Thr-Pen-Thr-NH₂, 0.5 μ M), but not by the κ -selective norbinaltorphimine (20 nM), when these antagonists were applied to the assay 1 min before or after the peptide. These results confirm the μ -selective bioactivity of the peptide.

The *in vivo* activity of the peptide as an analgesic was tested after intracerebroventricular (i.c.v.)¹⁷ and intrathecal¹⁸ injection to mice. Figure 2 (right) shows that after i.c.v. injection, the Phe-4 peptide, with a half-maximal effective dose (ED_{50}) of 4.7 nmol (95% confidence interval (CI) = 3.1–6.7) = 2.9 μ g (1.9–4.1 μ g) was at least as potent as morphine in producing analgesia (7.5 nmol (95% CI = 4.8–10.5) = 2.5 μ g (1.6–3.5)). Figure 2 (left) shows that prolonged analgesia can be induced by the Phe-4 peptide. About half the animals were still analgesic 1 hour after 10–20 μ g of the peptide. By contrast, the ED_{50} for the endogenous pentapeptide Met-enkephalin is 75 μ g and its effects last for less than 10 min¹⁹. The analgesia induced by 10 μ g of the Phe-4 peptide was reversed by a low dose of naloxone (1 mg kg⁻¹ injected

subcutaneously) and by pretreatment for 24 h with the irreversible μ -selective antagonist β -funaltrexamine (40 mg kg⁻¹), indicating that its action is mediated by μ receptors. The peptide was even more effective when given by intrathecal injection rather than by the i.c.v. route, with an ED_{50} of 0.8 ± 0.4 μ g. Thus, the Phe-4 peptide, an unmodified, all-natural amino-acid tetrapeptide, is a very potent analgesic.

We generated an antibody against the Phe-4 peptide which we used to develop a sensitive and specific radioimmunoassay that detects 1 pg peptide and shows no crossreactivity for over 40 opioid and non-opioid peptides and compounds. This radioimmunoassay enabled us to screen fractions from high-performance liquid chromatography of bovine cortical brain extracts and to isolate the peptide as described^{5,6}, with modifications (see below). Figure 3 shows the final purification step. The isolated material was subjected to six successive cycles of automated Edman degradation. The most prevalent phenylthiohydantoin (PTH)-amino acids were easily distinguishable from background amino acids, and the first four successive cycles yielded (in pmol): Tyr, 310; Pro, 154; Trp, 72; and Phe, 143. After the fourth cycle, no further PTH-amino acid was detected above background. The peptide structure was identified as Tyr-Pro-Trp-Phe-NH₂ and confirmed by comparison with synthetic Tyr-Pro-Trp-Phe-NH₂ and its free acid, Tyr-Pro-Trp-Phe-OH. The elution time of the synthetic peptide amide (21.9 min) was the same as that of the isolated peptide and was clearly separated from that of the peptide acid (25.9 min).

The only other amino acid that increased in the third cycle was Phe, with a value of 51 pmol, indicating the presence of the sequence

Table 1 Binding of agonist peptides to μ , δ and κ opiate receptors

	Receptor binding K_i (nM)			Binding selectivity	
	μ	δ	κ	δ/μ	κ/μ
DAMGO	0.34 $\pm$ 0.07	190 $\pm$ 16	1,300*	559	3,824
Endomorphin-1 (Tyr-Pro-Trp-Phe-NH ₂)	0.36 $\pm$ 0.08	1,506 $\pm$ 174	5,428 $\pm$ 474	4,183	15,077
Endomorphin-2 (Tyr-Pro-Phe-Phe-NH ₂)	0.69 $\pm$ 0.16	9,233 $\pm$ 201	5,240 $\pm$ 460	13,381	7,594
β -Endorphin	4.4 $\pm$ 0.41 2.1†	2.4†	96†	1.1	46
Met-enkephalin	5.9 $\pm$ 0.9 9.5†	0.91†	4,442†	0.01	468
Dynorphin	2.0 $\pm$ 0.5 0.73†	2.4†	0.12†	3.3	0.16
Tyr-W-MIF-1 (Tyr-Pro-Trp-Gly-NH ₂)*	70.9	15,520	22,300	219	314

Binding to μ (³H-DAMGO), δ (³H-pCl-DPDPE) or κ (³H-ethylketocyclazocine with 100 nM DAMGO and DPDPE to quench μ and δ binding) opiate receptors was determined as described in ref. 12. Values are means $\pm$ s.e.m. for 3 separate assays with 5–10 different concentrations of peptide.

* Values taken from ref. 12.

† For β -endorphin, Met-enkephalin and dynorphin, the means $\pm$ s.e.m. of 2 separate assays are shown together with values taken from ref. 3.

Table 2 Concentrations of endomorphin-1-like immunoreactivity in different regions of bovine brain

Region	Immunoreactivity (pmol per g tissue)
Thalamus	16.1
Hypothalamus	12.4
Striatum	10.2
Frontal cortex	8.3

Tyr-Pro-Phe-Phe-NH₂. This second peptide was synthesized and the retention time of the two synthetic peptides was similar (Tyr-Pro-Trp-Phe-NH₂, 21.9 min; Tyr-Pro-Phe-Phe-NH₂, 21.2 min). Again, the amidated peptide was clearly separated from that of the peptide acid (25.3 min). Like the two ligands for the δ receptor (Met- and Leu-enkephalin), the two new peptides differ by a single amino acid. The Trp-3 and Phe-3 in the new peptides represent the only two natural amino acids that maintain binding and activity in this position of the opioid pharmacophore^{20,21}, indicating the limit of similar natural sequences. From the 157 g of starting material, the isolated material from the Edman degradation steps was estimated to be ~200 ng, which consisted of ~75–85% Tyr-Pro-Trp-Phe-NH₂ and 15–25% Tyr-Pro-Phe-Phe-NH₂. At 2.1 pmol g⁻¹, this value indicates that minimal concentrations (after losses during isolation) may be below those of enkephalin, but similar to the less abundant opioids β -endorphin and dynorphin².

The second peptide showed a binding profile similar to the first, with about half the affinity for the μ site, but with a selectivity (>13,000-fold) for μ over δ sites of ~3-fold higher. Affinities at κ receptors were equally low for both peptides, making the first peptide about twice as selective (>15,000-fold) for μ over κ sites. The bioactivity of the second peptide approached that of the first *in vitro* (IC₅₀ in the guinea-pig ileum assay was 4 $\pm$ 0.4 nM) and *in vivo* ((ED₅₀ for analgesia after i.c.v. injection was 8.4 μ g (CI = 5.5–11.7)). Thus, both peptides have high affinity and selectivity as well as potent bioactivity.

The peptides isolated here differ structurally from previously known endogenous opioids in their N-terminal (Tyr-Pro) sequence, C-terminal amidation and tetrapeptide length. In addition, one of them is the first high-affinity ligand to contain Trp rather than Phe as the second aromatic moiety. Although other peptides with this feature have been reported^{22–24}, their affinity is below that of even the parent compound (Tyr-W-MIF-1)⁵, and the best known (hemorphin) is derived from enzymatic digests of blood²². β -Casomorphins²⁵ contain the Tyr-Pro-Phe sequence found in the second peptide, but contain the less active Pro in position 4, and are not found in neuronal tissue.

Because these are the first endogenous mammalian peptides that have a high affinity and a clear specificity for the opiate receptor preferred by morphine, we have named them endomorphin-1 (Tyr-Pro-Trp-Phe-NH₂) and endomorphin-2 (Tyr-Pro-Phe-Phe-NH₂).

Preliminary studies of the distribution of endomorphin-1 by radioimmunoassay indicate that it is found in thalamus, hypothalamus, cortex and striatum (Table 2). Thus, endomorphin-1-like immunoreactivity is present in several brain areas known to contain subregions of dense μ opiate receptors²⁶. We also anticipate subregions of high peptide density, however, that may not match those of dense receptor localization. Only ~7% of opiate receptors are estimated to be associated with synaptic junctions²⁷, suggesting that there is considerable extrasynaptic transmission by opioids^{28,29}. One advantage of selective, amidated peptides such as those described here would be to confer specificity and metabolic stability for such transmission.

In summary, we have discovered two new peptides in brain that have the highest affinity and specificity for the μ opiate receptor found so far in the mammalian nervous system. The peptides have potent μ -selective bioactivity, including analgesia. □

Methods

Isolation of endomorphins 1 and 2. Bovine frontal cortex was minced, boiled in 8 volumes of 0.08% Na₂S₂O₅ for 10 min, homogenized at 4 °C, adjusted to 25% acetonitrile in 1 M acetic acid, stirred for 4 h, and centrifuged at 29,000g for 20 min. Solid-phase extraction of the supernatants was achieved on Analytichem Mega Bond Elute C8 columns. Peptides were eluted with 70% acetonitrile, dried, and subjected to several sequential reversed-phase HPLC steps. Preparative separation on a Regis Prep-10D-60-ODS-FEC 25 $\times$ 2.1 cm column was done with an 8 ml min⁻¹ acetonitrile gradient increasing from 10 to 20, 30 and 80% over 10, 50 and 5 min; immunoreactivity eluted at fractions 41–45. Further purification of these fractions on a Vydac 201HS54 25 $\times$ 0.47 cm analytical column was at 1 ml min⁻¹ with acetonitrile increasing from 5 to 20, 25, 35, and 80% over 10, 30, 10 and 1 min. The immunoreactive fractions (23–25) were combined and chromatographed twice more using the same method and column; the final step is shown in Fig. 3.

Received 7 November 1996; accepted 7 February 1997.

1. Akil, H. *et al.* Endogenous opioids: Biology and function. *Annu. Rev. Neurosci.* **7**, 223–255 (1984).
2. Evans, C. J., Hammond, D. L. & Frederickson, R. C. A. in *The Opiate Receptors* (ed. Pasternak, G. W.) 23–71 (Humana, Clifton, New Jersey, 1988).
3. Corbett, A. D., Paterson, S. J. & Kostelitz, H. W. in *Opioids I (Handbook of Exp. Pharmacol.)* Vol. 104/1 (ed. Herz, A.) 645–679 (Springer, New York, 1993).
4. Reisine, T. & Pasternak, G. in *Goodman and Gilman's The Pharmacological Basis of Therapeutics* (eds Hardman, J. G. & Limbird, L. E.) 521–555 (McGraw-Hill, New York, 1996).
5. Erchegyi, J., Kastin, A. J. & Zadina, J. E. Isolation of a novel tetrapeptide with opiate and antipiate activity from human cortex: Tyr-Pro-Trp-Gly-NH₂ (Tyr-W-MIF-1). *Peptides* **13**, 623–631 (1992).
6. Hackler, L., Kastin, A. J., Erchegyi, J. & Zadina, J. E. Isolation of Tyr-W-MIF-1 from bovine hypothalamus. *Neuropeptides* **24**, 159–164 (1993).
7. Erchegyi, J. *et al.* Structure-activity relationships of analogs of the endogenous brain peptides Tyr-MIF-1 and Tyr-W-MIF-1. *Peptide Res.* **6**, 31–38 (1993).

8. Zadina, J. E., Kastin, A. J., Kenigs, V., Bruno, C. & Hackler, L. Prolonged analgesia after intracerebroventricular Tyr-W-MIF-1 (Tyr-Pro-Trp-Gly-NH₂). *Neurosci. Lett.* **155**, 220–222 (1993).
9. Gergen, K. A., Zadina, J. E., Kastin, A. J. & Paul, D. Intrathecal Tyr-W-MIF-1 produces potent, naloxone-reversible analgesia modulated by α_2 -adrenoceptors. *Eur. J. Pharmacol.* **298**, 235–239 (1996).
10. Zadina, J. E., Paul, D., Gergen, K. A., Ge, L.-J., Hackler, L. & Kastin, A. J. Binding of Tyr-W-MIF-1 and related peptides to μ_1 and μ_2 opiate receptors. *Neurosci. Lett.* **215**, 65–69 (1996).
11. Gergen, K. A., Zadina, J. E. & Paul, D. Analgesic effects of Tyr-W-MIF-1: A mixed μ_2 -opioid receptor agonist/ μ_1 -opioid receptor antagonist. *Eur. J. Pharmacol.* **316**, 33–38 (1996).
12. Zadina, J. E., Kastin, A. J., Ge, L.-J. & Hackler, L. Mu, delta and kappa opiate receptor binding of Tyr-MIF-1 and Tyr-W-MIF-1, its active fragments, and two potent analogs. *Life Sci.* **55**, 461–466 (1994).
13. Yamazaki, T., Ro, S., Goodman, M., Chung, N. N. & Schiller, P. W. A topochemical approach to explain morphine bioactivity. *J. Med. Chem.* **36**, 708–719 (1993).
14. Wang, J.-X., Bray, A. M., Dipasquale, A. J., Maeji, N. J. & Geysen, H. M. Systematic study of substance P analogs. I. Evaluation of peptides synthesized by the multipin method for quantitative receptor binding assay. *Int. J. Peptide Protein Res.* **42**, 384–391 (1993).
15. Kyte, J. & Doolittle, R. F. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157**, 105–132 (1982).
16. Zadina, J. E., Kastin, A. J., Kersh, D. & Wyatt, A. Tyr-MIF-1 and hemorphin can act as opiate agonists as well as antagonists in the guinea pig ileum. *Life Sci.* **51**, 869–885 (1992).
17. Haley, T. J. & McCormick, W. G. Pharmacological effects produced by intracerebral injection of drugs in the conscious mouse. *Br. J. Pharmacol. Chemother.* **12**, 12–15 (1957).
18. Hylden, J. L. K. & Wilcox, G. L. Intrathecal morphine in mice: A new technique. *Eur. J. Pharmacol.* **67**, 313–316 (1980).
19. Buscher, H. H. *et al.* Evidence for analgesic activity of enkephalin in the mouse. *Nature* **261**, 423–425 (1976).
20. Schiller, P. W., Yam, C. F. & Prossman, J. Synthesis, opiate receptor affinity, and conformational parameters of [4-Tryptophan] enkephalin analogues. *J. Med. Chem.* **21**, 1110–1116 (1978).
21. Lenz, G. R., Evans, S. M., Walters, D. E. & Hopfinger, A. J. *Opiates* (Academic, Orlando, Florida, 1986).
22. Brantl, V. *et al.* Novel opioid peptides derived from hemoglobin: hemorphins. *Eur. J. Pharmacol.* **125**, 309–310 (1986).
23. Erchevgyi, J., Kastin, A. J., Zadina, J. E. & Qiu, X.-D. Isolation of a heptapeptide Val-Val-Tyr-Pro-Tyr-Thr-Gln (valorphin) with some opiate activity. *Int. J. Peptide Protein Res.* **39**, 477–484 (1992).
24. Glamsta, E.-L. *et al.* Isolation and characterization of a hemoglobin-derived opioid peptide from the human pituitary gland. *Regul. Pept.* **34**, 169–179 (1991).
25. Brantl, V. Novel opioid peptides derived from human β -casein: Human β -casomorphins. *Eur. J. Pharmacol.* **106**, 213–214 (1984).
26. Mansour, A., Fox, C. A., Akil, H. & Watson, S. J. Opioid-receptor mRNA expression in the rat CNS: Anatomical and functional implications. *Trends Neurosci.* **18**, 22–29 (1995).
27. Hamel, E. & Beaudet, A. Electron microscopic autoradiographic localization of opioid receptors in rat neostriatum. *Nature* **312**, 155–157 (1984).
28. Herkenham, M. In *Volume Transmission in the Brain: Novel Mechanisms for Neural Transmission* (eds Fuxe, K. & Agnati, L. F.) 63–87 (Raven, New York, 1991).
29. Elde, R. *et al.* in *Diversity of Interacting Receptors* (eds Abood, L. G. & Lajtha, A.) 390–404 (*Acad. Sci.*, New York, 1995).
30. Yoburn, B. C., Cohen, A., Umans, J. G., Ling, G. F. & Inturrisi, C. E. The graded and quantal nature of opioid analgesia in the rat tailflick assay. *Brain Res.* **331**, 327–336 (1985).

Acknowledgements. We thank X. Zhang and G. Wilde for technical assistance, W. A. Banks for suggestions on the manuscript, and NIDA for DAMGO, DPDPE and ³H-EKC. Supported by the VA.

Correspondence and requests for materials should be addressed to J.E.Z.

Protease-activated receptor 3 is a second thrombin receptor in humans

Hiroaki Ishihara*, Andrew J. Connolly†, Dewan Zeng*, Mark L. Kahn*, Yao Wu Zheng*, Courtney Timmons*, Tracy Tram* & Shaun R. Coughlin*‡

* Cardiovascular Research Institute and Daiichi Research Center, Departments of † Pathology and ‡ Medicine, University of California, San Francisco, HSW-831, 505 Parnassus Avenue, San Francisco, California 94143-0130, USA

Thrombin is a coagulation protease that activates platelets, leukocytes, endothelial and mesenchymal cells at sites of vascular injury, acting partly through an unusual proteolytically activated G-protein-coupled receptor^{1–3}. Knockout of the gene encoding this receptor provided definitive evidence for a second thrombin receptor in mouse platelets and for tissue-specific roles for different thrombin receptors⁴. We now report the cloning and characterization of a new human thrombin receptor, designated protease-activated receptor 3 (PAR3). PAR3 can mediate thrombin-triggered phosphoinositide hydrolysis and is expressed in a variety of tissues, including human bone marrow and mouse megakaryocytes, making it a candidate for the sought-after

second platelet thrombin receptor. PAR3 provides a new tool for understanding thrombin signalling and a possible target for therapeutics designed selectively to block thrombotic, inflammatory and proliferative responses to thrombin.

A polymerase chain reaction (PCR)-based strategy yielded a new human complementary DNA encoding a putative G-protein-coupled receptor with 27% amino-acid sequence similarity to the cloned human thrombin receptor³ (henceforth called PAR1) and 28% similarity to PAR2 (Fig. 1a). PAR2, a possible trypsin receptor, is the only other known member of the protease-activated receptor family⁵. The amino-terminal exodomain of the new receptor, designated PAR3, contained a possible thrombin cleavage site at residues K38/T39, followed by a sequence strikingly identical to a thrombin-binding sequence in the leech anticoagulant hirudin⁶ (Fig. 1b). Analogous sequences in PAR1 mediate recognition and efficient cleavage by thrombin^{7–10} (Fig. 1b); this cleavage unmasks a new amino terminus which serves as a tethered peptide ligand, binding intramolecularly to the body of PAR1 to effect transmembrane signalling^{3,11,12}. These observations indicated that PAR3 was a new thrombin receptor that should be activated by cleavage of the K38/T39 peptide bond.

We confirmed that PAR3 was a thrombin substrate in a reaction in which 20 nM thrombin cleaved 80% of PAR3 expressed on the surface of Cos 7 cells within 5 min, which is comparable to PAR1 cleavage (Fig. 2). PAR3 cleavage was prevented by substitution of proline for threonine at position 39 (Fig. 2), the P1' residue in the putative K38/T39 cleavage site (Fig. 1). To confirm the location of the thrombin cleavage site on PAR3, the amino-terminal exodomain of PAR3 (residues 21–94) was expressed in *Escherichia coli* as a soluble polypeptide and cleaved in solution. Even when incubated with 50 nM thrombin for 1 hour at 37°C, only two cleavage products were detected on SDS-PAGE. Their size was consistent with cleavage at the K38/T39 peptide bond, and amino-terminal sequencing revealed only the original amino terminus and a single new amino terminus with the sequence TFRGA (Fig. 1). Thus thrombin selectively cleaves the peptide bond between K38 and T39 in PAR3's amino-terminal exodomain.

Does PAR3 cleavage by thrombin trigger signalling? Cos 7 cells transfected with PAR3 cDNA gave robust thrombin-stimulated phosphoinositide hydrolysis (Fig. 3a). Co-transfection with α_{16} , a G-protein α -subunit expressed in haematopoietic cell lines¹³, caused a 50–100% increase in the maximal PAR3-mediated response to thrombin in these cells (Fig. 3a). The half-maximal effective concentration (EC₅₀) for thrombin signalling by PAR3 in this system was ~0.2 nM, about twice that obtained with PAR1 and well within physiological thrombin concentrations (Fig. 3b). Mutation of the K38/T39 site to prevent receptor cleavage (Fig. 2) ablated PAR3 signalling (Fig. 3a), and thrombin rendered proteolytically inactive by the inhibitor PPACK¹⁴ caused no signalling in PAR3-transfected cells at concentrations as high as 1 μ M (not shown). Thus PAR3 mediates thrombin signalling and PAR3's K38/T39 peptide bond must be cleaved for activation of the receptor.

Several observations suggest that PAR3, like PAR1, uses thrombin's fibrinogen-binding exosite for receptor recognition^{7–10}. γ -thrombin, which is defective in its fibrinogen-binding exosite¹⁵, was 100 times less potent than α -thrombin (Fig. 3b). Similarly, incubation of α -thrombin with the fibrinogen-binding exosite blocker hirugen¹⁶ right-shifted the dose-response curve by a factor of 100 (data not shown). Alanine substitutions at F48 and E49 in PAR3's hirudin-like sequence, residues predicted to dock with thrombin's fibrinogen-binding exosite by analogy with PAR1 (refs 7, 9) and hirudin⁶ (Fig. 1), attenuated thrombin cleavage (Fig. 2) and right-shifted PAR3's concentration response to thrombin tenfold (data not shown).

How specific is PAR3 for thrombin compared with other proteases? We compared protease-triggered mobilization of radioactive calcium (⁴⁵Ca), which reflects phosphoinositide hydrolysis in

Xenopus oocytes expressing PAR3 or PAR1 (ref. 7). Thrombin caused strong responses in PAR3-expressing oocytes, whereas other arginine/lysine-specific serine proteases (factor Xa, trypsin, factor VIIa, tissue plasminogen activator or plasmin) had little or no activity (Fig. 3c). Chymotrypsin, elastase and cathepsin G were also inactive. Oocytes expressing PAR1 responded even more strongly to thrombin and were activated to some extent by preparations of factor Xa, plasmin or trypsin. Thus PAR3 is at least as specific for thrombin as the bona fide thrombin receptor PAR1.

Synthetic peptides that mimic the tethered ligands unmasked by thrombin cleavage of PAR1 or trypsin cleavage of PAR2 activate those receptors independently of receptor cleavage^{3,5}. Surprisingly, peptides mimicking the putative tethered ligand of PAR3 (TFRGAP and TFRGAPPNS) or the PAR1- and PAR2-activating peptides SFLLRN and SLIGRL, showed little or no activity at PAR3 at concentrations as high as 100 μ M (data not shown). PAR3's putative tethered ligand domain is required for signalling in that substitution of alanine for Phe 40 in PAR3 (Fig. 1b) yielded a receptor that failed to signal upon thrombin cleavage (Figs 2, 3). F40 in PAR3 is analogous to the critical F43 in PAR1's tethered ligand¹⁷ (Fig. 1b). Cleavage of the K38/T39 peptide bond in PAR3 may trigger its activation by the same tethered-ligand mechanism^{3,11} that

is used by PAR1 and 2; failure of synthetic tethered ligand to activate PAR3 may be explained by a different avidity or structural context. We cannot yet exclude alternative mechanisms of activation: for example, PAR3 cleavage may release a constraint that maintains the receptor in an 'off' state.

Northern blot analysis revealed two PAR3 messenger RNA species migrating at 3.4 kilobases (kb) and 1.8 kb in human and at 2.8 kb and 1.8 kb in mouse (Fig. 4a). Hybridization with probes for coding as opposed to 3' untranslated regions of mouse PAR3 cDNA and sequencing of 3' rapid amplification of cDNA ends (RACE) products showed that the smaller mouse mRNA resulted from use of an alternative polyadenylation signal (ATTAAA) 340 nucleotides 3' of the stop codon. The human cDNA sequence also contained a potential alternative polyadenylation signal (ACTAAA) 260 nucleotides 3' of the stop codon. In humans, PAR3 mRNAs are expressed in bone marrow and a variety of other tissues (Fig. 4a), although it is not known which cell types are responsible for this expression. In mouse, PAR3 mRNA was detected in spleen, with a less intense signal in lung and brain (Fig. 4b). *In situ* hybridization revealed that PAR3 expression was high in megakaryocytes in mouse spleen and bone marrow (Fig. 4c, and data not shown) but was low in other mouse tissues under our exposure conditions (Fig. 4c).

a

```

hPAR3- 1 MKA LIFAAAGLLLLLP TFCQSGMENDTNNLAKP TLPIK/TFRGAPPN SFEFFPFSALEGWTGATITVKIKC PEESASHLHVKNATMG
hPAR1- 1 MGPRR LLLVAACFSLCGP LLSARTRARRPESKATNATLDPR/SFLLRNPNDKYEPPFWEDEEKNESGLTEYRLVSINKSSPLQKQPAFISEDASG
hPAR2- 1 MRSPSAWLLGAAILLA ASLSCSGTIQG TNRSSKGR/SLIGKVDGTSHTVGKGVTV ETVFSVDFEFSAS

|-----TM1-----| |-----TM2-----| |-----TM3-----|
hPAR3- 87 YLTSSLSTKLIPAIYLLVFVGVANAVTLWMLFFRTR SICTTVEYTNLAIAIDFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTIVFYGNMYCSILLACISINRYLAI
hPAR1- 95 YLTSSLSTKLIPAIYLLVFVGVANAVTLWMLFFRTR SICTTVEYTNLAIAIDFLFCVTLPFKIAYHLNGNNWVFGEVLCRATTIVFYGNMYCSILLACISINRYLAI
hPAR2- 68 VLTGKLTITVPLPIVYITVVFVGLPSNGMALWVFLRTKKKHPAVIYMANLALADLLSVIWFPLKIAHYHGNWYIYGEALCNVLIGFFYGNMYCSILEMTCLSVQRYWVI

|-----TM4-----| |-----TM5-----|
hPAR3- 196 VHPFTYRGLPKHYALVTCGLVWATVFLYMLPFFILKQEYLLVQPDITTCCHDVHNTCESSPPQLYYFISLAFFGFLIPFVLIICYAAIIRTUNA YDHRWLWYV
hPAR1- 205 VYPMQSLSWRTLGRASFTCLAIWALAIAGVPLVLKQETIQVPLNITTCCHDVHNTLLEGG YYAYYFSAFSAVFFVPLIISTVCYVSIIRCLSSSAVANRSKK SRAL
hPAR2- 178 VNPMSHSRKKANIAIGI SLAIWLLILLVTIPLVVKQITIFIPALNITTCCHDVLPQLLVGD MFNYFLSLAIGVFLFPAPLTASAYVLMIRMLRSSAMDENSEKKRRKAI

|-----TM6-----| |-----TM7-----|
hPAR3- 301 KASLLILVITICFAPSNIILIIHHANYYNNT DGLYFIYLIALCLGSLNSCLDPFLYFLMSKTRNHSYAYLTK
hPAR1- 313 FLAAVFCIFICFGPTNVLLIAHYSFLSHTSTTEAAYFAYLLCVCSISSCIDPLIYYASSECQRYVYSILCKESSDPSSYNSSGQLMASKMDTCSNLLNNSIYKLLT
hPAR2- 287 KLIVTVLAMYLICFTPSNLLLVVHY FLIKSQGQSHVYALYIVALCLSTLNSCIDPFVYFVSHDFRDHAKNALLCRSVRTVKQVQVSLTSKKHSRKKSSSYSSSTTVKTSY

```

b

```

Hirudin C-tail
hPAR3- 34-57 . . . TLPIK / TFRGAPPN SFEFFPFSALE . . .
hPAR1- 37-61 . . . TLDPR / SFLLRNPNDKYEPPFWEDEEK . . .
hPAR2- 32-56 . . . SSKGR / SLIGKVDGTSHTVGKGVTV . . .

```

Cleavage site

Extracellular

Cytoplasmic

Figure 1 Amino-acid sequence of human PAR3. **a**, Alignment with PAR1 and PAR2. Predicted transmembrane (TM) domains are overlined. Sequences corresponding to the degenerate PCR primers are underlined. **b**, Features of PAR amino-terminal exodomains. Cleavage sites are indicated; tethered ligand domains of PAR1 and PAR2 and putative tethered ligand domain of PAR3 are underlined. Also underlined is PAR3's hirudin-like domain (FEEFP). Note the similarity to the FEEIP and YEPFW sequences in hirudin and PAR1 that bind to thrombin's fibrinogen-binding exosite⁸⁻¹⁰.

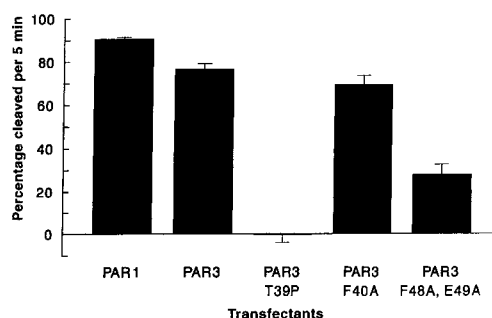


Figure 2 Thrombin cleaves PAR3. Wild-type or mutant human PAR1 or PAR3 cDNAs encoding receptors displaying a Flag epitope (see Methods) amino to their thrombin cleavage sites were transiently expressed in Cos7 cells. The T39P mutation renders the predicted thrombin-cleavage site uncleavable, F40A hobbles PAR3's putative tethered ligand domain, and F48A, E49A mutates two residues predicted to contribute to thrombin binding (Fig. 1b). Surface expression levels, assessed as transfection-dependent binding of M1 anti-Flag antibody to the cell surface, were similar for all constructs examined. Cells were incubated for 5 min at 37 °C in the presence or absence of 20 nM thrombin. Receptor cleavage was detected as loss of the Flag epitope from the cell surface (expressed as a percentage of M1-antibody-specific binding lost upon thrombin treatment). Data shown are mean $\pm$ s.e.m. ($n = 3$). This experiment was replicated three times.

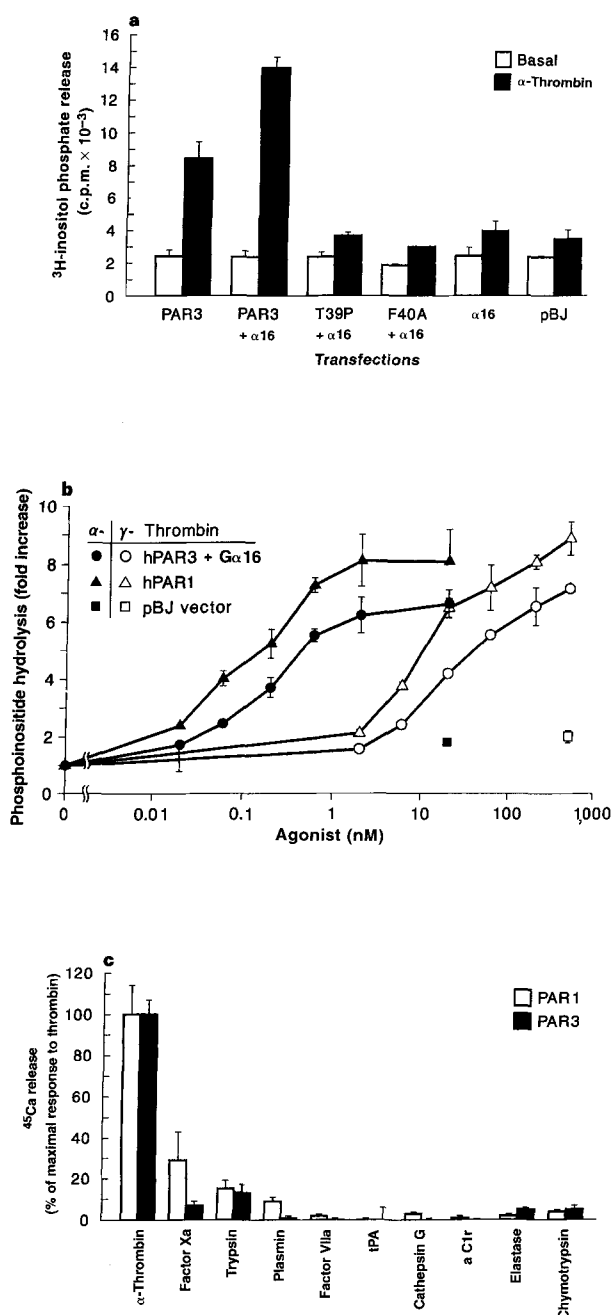


Figure 3 PAR3 signalling. **a**, Wild-type and mutant PAR3 signalling in Cos 7 cells. Epitope-tagged human PAR1, PAR3, or the indicated mutants were transiently expressed in Cos 7 cells with or without $G\alpha 16$. Expression was similar for wild-type and mutant receptors with or without $G\alpha 16$ coexpression. Data shown are mean $\pm$ s.e.m. ($n = 3$) 3H -inositol phosphates released per 120 min in the presence or absence of 20 nM α -thrombin²⁰. Similar results were obtained when the incubation time was 15 min. **b**, Concentration responses. Phosphoinositide hydrolysis in response to the indicated concentration of α - or γ -thrombin was measured in Cos 7 cells transfected with PAR1 or PAR3 plus $G\alpha 16$, as in **a**. Data are mean $\pm$ s.e.m. ($n = 3$). **c**, Specificity of PAR3 and PAR1. Protease-triggered ^{45}Ca release per 10 min was measured in *Xenopus* oocytes expressing human PAR1 or PAR3. Proteases were at 20 nM except trypsin (0.5 nM). Data are mean $\pm$ s.e.m. ($n = 2$) and are expressed as a percentage of the maximum response to thrombin. In the experiment shown, 20 nM thrombin caused 24- and eightfold increases in ^{45}Ca release in PAR1- and PAR3-expressing oocytes, respectively, and surface expression of PAR1 was twice that of PAR3. No responses were seen in uninjected oocytes. Experiments in **a**, **b** and **c** were replicated three times.

Mouse platelets use a thrombin receptor that is distinct from PAR1 (ref. 4). The PAR3 expression found in mouse megakaryocytes indicates that PAR3 may mediate thrombin responses in mouse platelets. Determination of the signalling properties of mouse PAR3 and knockout of the PAR3 gene will test this idea. Unlike mouse platelets, human platelets use PAR1 (refs 3, 17–19) but thrombin's greater efficiency compared with PAR1-activating peptide at triggering some responses in human platelets suggests that they may use PAR1 and a second thrombin receptor^{20–23}.

Defining the signalling pathways and relative roles for PAR1 and PAR3 in human platelets, endothelium, leukocytes, mesenchymal and other cells will be important. The availability of different thrombin receptors with tissue-specific roles may allow more effective or selective disruption of thrombin signalling and help in development of pharmaceuticals aimed at interrupting thrombin's thrombotic, inflammatory and proliferative actions.

Note added in proof: A partial sequence of mouse PAR3 was recently deposited in the WashU-HHMI EST Project database. Accession no. mt07f02.r1.

Methods

Cloning of PAR3. Rat platelets were included among many RNA sources tested because they are a more abundant source of RNA than mouse platelets and, like mouse platelets, they do not respond to PAR1-agonist peptides (refs 4, 24, 25, and data not shown). Total RNA was prepared from rat platelets using Trizol reagent (Gibco BRL). cDNA was prepared using random hexamer primers and Superscript reverse transcriptase (Gibco BRL), then amplified using degenerate PCR primers. 11 'forward' and 13 'reverse' degenerate primers corresponding to a variety of sites relatively conserved among PAR1, PAR2, peptide, and glycoprotein hormone GPCRs²⁶ were tested in different pairwise combinations. Amplification of rat platelet cDNA with the primers 5'-GTI TAC ATG CTI (A/C)AC (C/T)TI GCI (A/C/G/T)TI GC(A/C/G/T) GA-3' and 5'-GGA TAI ACI GCI A(A/G/T)(A/G) (A/T)AI C(G/T)(A/C/G/T) TC-3' yielded PAR3. Primers were used at 5 μ M in 20 mM Tris-HCl (pH 8.4), 50 mM KCl, 1.5 mM $MgCl_2$, 0.2 mM dNTP, with 50 U ml⁻¹ *Taq* polymerase. Temperature cycles were: 94 °C for 4 min; 30 cycles of 94 °C for 45 s, 39 °C for 60 s, 72 °C for 90 s; then 72 °C for 7 min. PCR products were subcloned using a TA cloning kit (Invitrogen). Clones with inserts of ~200 bp were grouped by restriction analysis and analysed by nucleic acid sequencing. One sequence predicted a novel G-protein-coupled receptor highly related to PAR1 and PAR2. This rat sequence was used to obtain mouse and human cDNA clones by a combination of standard PCR and hybridization techniques. The human PAR3 cDNA used for these studies was from a λ gt10 small intestine cDNA library (Clontech). The mouse cDNA used for *in situ* hybridization and northern blot analysis was from a spleen library in Lambda ZAP (Stratagene).

PAR1 and PAR3 functional studies. cDNA for epitope-tagged PAR3 analogous to Flag-epitope-tagged PAR1²⁷ was constructed to encode an N terminus with the sequence MDSKGSSQKGSRLLLLVSNLLCQGVVS/DYKDDDDVE-TF representing the prolactin signal peptide, the putative signal peptidase site (/), the Flag epitope DYKDDDD and junction VE fused to amino acid 17 in PAR3. Receptor cDNAs were subcloned into the mammalian expression vector pBJ1 or pFROG³. For receptor cleavage, Cos 7 cells were transfected using DEAE-dextran and thrombin-dependent loss of M1 antibody (Kodak) binding to the cell-surface Flag epitope was followed^{10,27}. Over 95% of M1 antibody binding was transfection-dependent in this system. For identification of the cleavage site, PAR3 N-terminal exodomain residues 21–94 sandwiched between a translational start and hexahistidine tag (that is, MG, PAR3 21–94, VEHHHHHH) were expressed as a soluble polypeptide in *E. coli*, purified, and analysed before and after thrombin cleavage, as described for the analogous region in PAR (ref. 10). For signalling studies^{3,28,29}, *Xenopus* oocytes were microinjected with 12.5–25 ng receptor cRNA/oocyte and Cos 7 cells were transfected with 5 μ g receptor cDNA in pBJ1 and/or 5 μ g $G\alpha 16$ cDNA in pCDNA1 (Invitrogen) per 10-cm plate.

PAR3 expression in mouse and human tissues. Northern blots with ~2 μ g mRNA loaded per lane (Clontech) were hybridized at high stringency with ^{32}P -labelled probes for human or mouse PAR3. Probes were generated by random priming (Prime-It II kit; Stratagene) of PCR-amplified DNA fragments

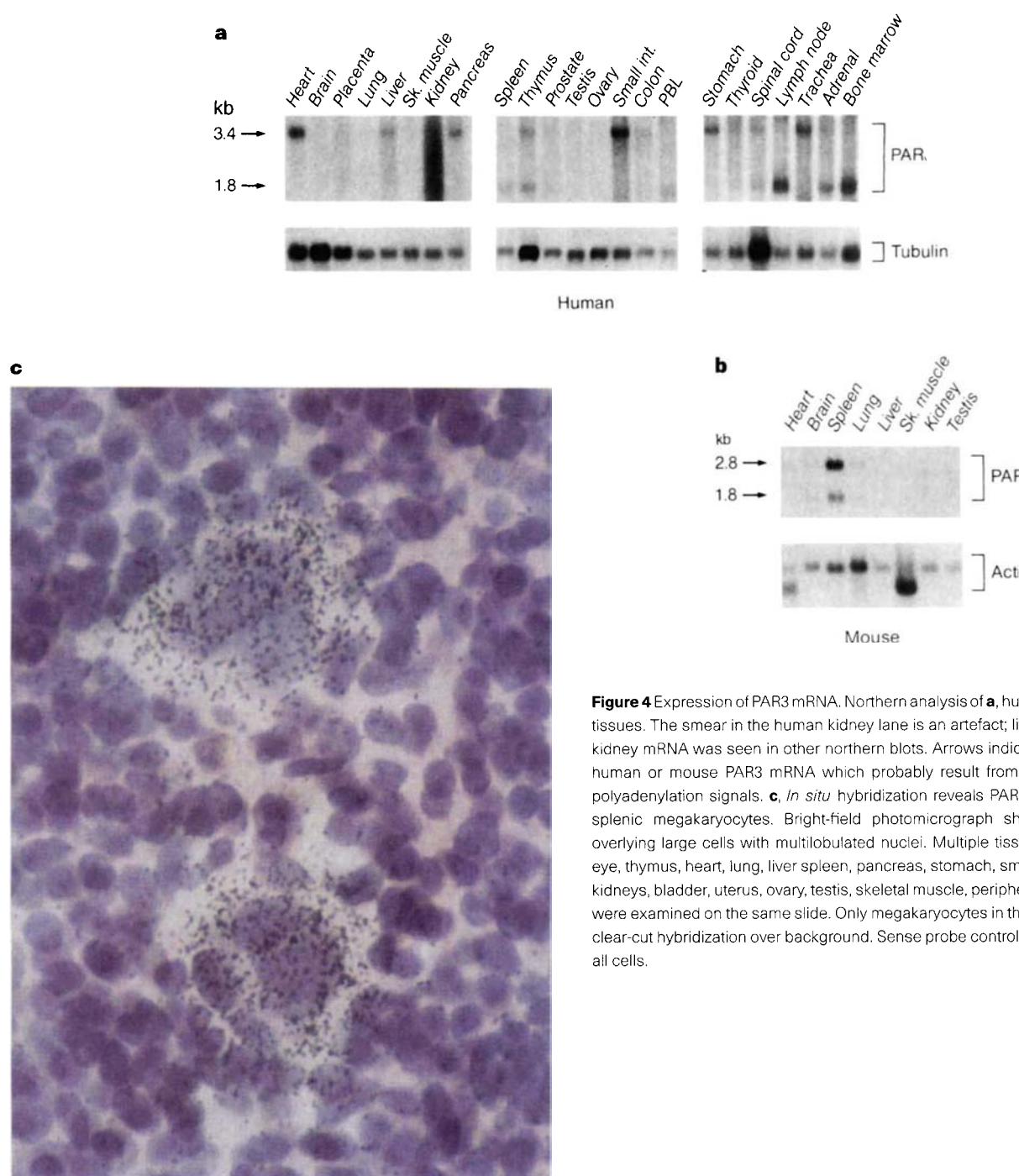


Figure 4 Expression of PAR3 mRNA. Northern analysis of **a**, human, and **b**, mouse tissues. The smear in the human kidney lane is an artefact; little hybridization to kidney mRNA was seen in other northern blots. Arrows indicate two species of human or mouse PAR3 mRNA which probably result from use of alternative polyadenylation signals. **c**, *In situ* hybridization reveals PAR3 mRNA in mouse splenic megakaryocytes. Bright-field photomicrograph shows silver grains overlying large cells with multilobulated nuclei. Multiple tissues (spleen, brain, eye, thymus, heart, lung, liver spleen, pancreas, stomach, small intestine, colon, kidneys, bladder, uterus, ovary, testis, skeletal muscle, peripheral nerve and skin) were examined on the same slide. Only megakaryocytes in the spleen displayed clear-cut hybridization over background. Sense probe controls were negative for all cells.

corresponding to nucleotides 459–860 of human and 414–564 of mouse PAR3 cDNAs (A of start ATG, represents the first nucleotide). Hybridization for human tubulin or mouse β -actin controlled for lane loading. For *in situ* hybridization, anaesthetized adult C57BL/6 mice were perfusion-fixed with 4% paraformaldehyde. Organs were dissected free, immersion-fixed for 4 h in 4% paraformaldehyde, then embedded in paraffin. 5- μ m sections were hybridized with a sense or antisense 35 S-riboprobe transcribed *in vitro* from mouse PAR3 cDNA subcloned into the *Eco*RI site of pBluescript II Sk⁺. Hybridization, washing and development conditions were as described for mouse PAR1 (ref. 30). A 4-week exposure result is shown.

Received 4 November 1996; accepted 5 February 1997.

1. Fenton, J. W. Regulation of thrombin generation and functions. *Sem. Hemostas. Thrombos.* **14**, 234–240 (1988).
2. Coughlin, S. R. Thrombin receptor function and cardiovascular disease. *Trends Cardiovas. Med.* **4**, 77–83 (1993).

3. Vu, T.-K. H., Hung, D. T., Wheaton, V. I. & Coughlin, S. R. Molecular cloning of a functional thrombin receptor reveals a novel proteolytic mechanism of receptor activation. *Cell* **64**, 1057–1068 (1991).
4. Connolly, A., Ishihara, H., Kahn, M., Farese, R. V. Jr & Coughlin, S. R. Role of the thrombin receptor in development and evidence for a second receptor. *Nature* **381**, 516–519 (1996).
5. Nystedt, S., Emilsson, K., Wahlestedt, C. & Sundelin, J. Molecular cloning of a potential novel proteinase activated receptor. *Proc. Natl Acad. Sci. USA* **91**, 9208–9212 (1994).
6. Rydel, T. J. *et al.* The structure of a complex of recombinant hirudin and human α -thrombin. *Science* **249**, 277–280 (1990).
7. Vu, T.-K. H., Wheaton, V. I., Hung, D. T. & Coughlin, S. R. Domains specifying thrombin-receptor interaction. *Nature* **353**, 674–677 (1991).
8. Liu, L., Vu, T.-K. H., Esmon, C. T. & Coughlin, S. R. The region of the thrombin receptor resembling hirudin binds to thrombin and alters enzyme specificity. *J. Biol. Chem.* **266**, 16977–16980 (1991).
9. Mathews, I. I. *et al.* Crystallographic structures of thrombin complexed with thrombin receptor peptides: existence of expected and novel binding modes. *Biochemistry* **33**, 3266–3279 (1994).
10. Ishii, K., Gerszten, R., Zheng, Y.-W., Turck, C. W. & Coughlin, S. R. Determinants of thrombin receptor cleavage: Receptor domains involved, specificity, and role of the P3 aspartate. *J. Biol. Chem.* **270**, 16435–16440 (1995).
11. Chen, J., Ishii, M., Wang, L., Ishii, K. & Coughlin, S. R. Thrombin receptor activation: Confirmation of the intramolecular tethered ligand hypothesis and discovery of an alternative intermolecular ligand mode. *J. Biol. Chem.* **269**, 16041–16045 (1994).

12. Gerszten, R. E. *et al.* The thrombin receptor's specificity for agonist peptide is defined by its extracellular surface. *Nature* **368**, 648–651 (1994).
13. Amatruda, T. T., Steele, D. A., Slepak, V. Z. & Simon, M. I. α subunit specifically expressed in hematopoietic cells. *J. Biol. Chem.* **266**, 5587–5591 (1991).
14. Kettner, C. & Shaw, E. D. *p*-phe-pro-arg-CH₂Cl: a selective affinity label for thrombin. *Thrombosis Res.* **14**, 969–973 (1979).
15. Rydel, T. J. *et al.* Crystallographic structure of human γ -thrombin. *J. Biol. Chem.* **269**, 22000–22006 (1994).
16. Skrzypczak, J. E. *et al.* Structure of the hirugen and hirulog 1 complexes of α -thrombin. *J. Mol. Biol.* **221**, 1379–1393 (1991).
17. Scarborough, R. M. *et al.* Tethered ligand agonist peptides: structural requirements for thrombin receptor activation reveal mechanism of proteolytic unmasking of agonist function. *J. Biol. Chem.* **267**, 13146–13149 (1992).
18. Hung, D. T., Vu, T.-K. H., Wheaton, V. L., Ishii, K. & Coughlin, S. R. The cloned platelet thrombin receptor is necessary for thrombin-induced platelet activation. Blocking antiserum to the thrombin receptor's hirudin-like domain. *J. Clin. Invest.* **89**, 1350–1353 (1992).
19. Brass, L. F. *et al.* Structure and function of the human platelet thrombin receptor. Studies using monoclonal antibodies directed against a defined domain within the receptor N terminus. *J. Biol. Chem.* **267**, 13795–13798 (1992).
20. Kinlough Rathbone, R., Perry, D. W., Guccione, M. A., Rand, M. L. & Packham, M. A. Degranulation of human platelets by the thrombin receptor peptide SFLLRN: comparison with degranulation by thrombin. *Thromb. Haemost.* **70**, 1019–1023 (1993).
21. Goodwin, C. A. *et al.* Thrombin receptor activating peptide does not stimulate platelet procoagulant activity. *Biochem. Biophys. Res. Commun.* **202**, 321–327 (1994).
22. Iau, L. F., Pumiaglia, K., Cote, Y. P. & Feinstein, M. B. Thrombin-receptor agonist peptides, in contrast to thrombin itself, are not full agonists for activation and signal transduction in human platelets in the absence of platelet-derived secondary mediators. *Biochem. J.* **303**, 391–400 (1994).
23. Kramer, R. M. *et al.* Differential activation of cytosolic phospholipase A2 (cPLA2) by thrombin and thrombin receptor agonist peptide in human platelets. evidence for activation of cPLA2 independent of the mitogen-activated protein kinases ERK1/2. *J. Biol. Chem.* **270**, 14816–14823 (1995).
24. Connolly, T. M. *et al.* Species variability in platelet and other cellular responsiveness to thrombin receptor-derived peptides. *Thromb. Haemost.* **72**, 627–633 (1994).
25. Derian, C. K., Santulli, R. J., Tomko, K. A., Haertlein, B. J. & Andrade-Gordon, P. Species differences in platelet responses to thrombin and SFLLRN. Receptor-mediated calcium mobilization and aggregation and regulation by protein kinases. *Thrombosis Res.* **6**, 505–519 (1995).
26. Probst, W. C., Snyder, L. A., Schuster, D. J., Brosius, J. & Sealfon, S. C. Sequence alignment of the G-protein coupled receptor superfamily. *DNA and Cell* **11**, 1–20 (1992).
27. Ishii, K., Hein, L., Kobilka, B. & Coughlin, S. R. Kinetics of thrombin receptor cleavage on intact cells: relation to signaling. *J. Biol. Chem.* **268**, 9780–9786 (1993).
28. Williams, J. A., McChesney, D. J., Calayag, M. C., Lingappa, V. R. & Logsdon, C. D. Expression of receptors for cholecystokinin and other Ca^{2+} -mobilizing hormones in *Xenopus* oocytes. *Proc. Natl Acad. Sci. USA* **85**, 4939–4943 (1988).
29. Nanevicz, T., Wang, L., Chen, M., Ishii, M. & Coughlin, S. R. Activating mutations in the thrombin receptor's agonist recognition domain. Mutations in a G protein-coupled receptor's extracellular domain cause transmembrane signaling. *J. Biol. Chem.* **271**, 702–706 (1996).
30. Soifer, S. J., Peters, K. G., O'Keefe, J. & Coughlin, S. R. Disparate temporal expression of the prothrombin and thrombin receptor genes during mouse development. *Am. J. Pathol.* **144**, 60–69 (1993).

Acknowledgements. We thank C. Turck for peptide sequencing, J. Fenton II for α -thrombin and its derivatives, M. Simon for α -thrombin cDNA, M. Davis for pBJ1, and H. Bourne for critical reading of this manuscript.

Correspondence and requests for materials should be addressed to S.R.C. (e-mail: shaun_coughlin.src@quickmail.ucsf.edu). Sequences of mouse and human cDNAs for PAR3 have been deposited in GenBank, accession numbers U92971 and U92972.

A second serine protease associated with mannan-binding lectin that activates complement

Steffen Thiel, Thomas Vorup-Jensen, Cordula M. Stover*, Wilhelm Schwaebler*, Steen B. Laursen, Knud Poulsen, Anthony C. Willis†, Paul Eggleton‡, Søren Hansen, Uffe Holmskov‡, Kenneth B. M. Reid† & Jens C. Jensenius

Department of Medical Microbiology and Immunology, University of Aarhus, DK 8000, Aarhus, Denmark

* Department of Microbiology and Immunology, University of Leicester, Leicester LE1 9HN, UK

† MRC Immunochemistry Unit, University of Oxford, Oxford OX1 3QU, UK

‡ Department of Medical Microbiology, University of Odense, Odense DK 5000, Denmark

The complement system comprises a complex array of enzymes and non-enzymatic proteins that is essential for the operation of the innate as well as the adaptive immune defence¹. The complement system can be activated in three ways: by the classical pathway which is initiated by antibody-antigen complexes, by

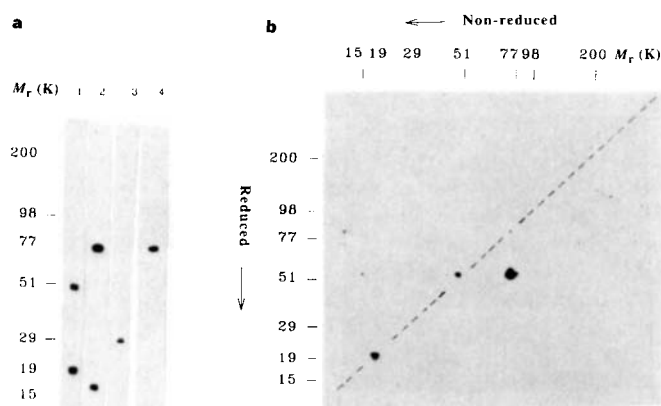


Figure 1 Western blotting of human plasma proteins purified by sugar affinity chromatography. One- and two-dimensional SDS-PAGE of this preparation was followed by blotting onto polyvinylidene difluoride (PVDF) membranes and development with various antibodies. **a**, One-dimensional electrophoresis; lanes 1 and 3 where run under reducing conditions and lanes 2 and 4 under non-reducing conditions. Lanes 1 and 2 were developed with anti-N¹MASP-2 antibody, lanes 3 and 4 with anti-C¹MASP-2 antibody. **b**, SDS-PAGE in two dimensions; the first dimension was run under non-reducing conditions. The lane was cut out, incubated in sample buffer containing DTT, placed on top of another SDS-polyacrylamide gel and electrophoresed; the gel was then blotted and the blot developed with anti-N¹MASP-2 antibody. The positions of M_r markers are indicated.

the alternative pathway initiated by certain structures on microbial surfaces, and by an antibody-independent pathway² that is initiated by the binding of mannan-binding lectin (MBL; first described as mannan-binding protein³) to carbohydrates. MBL is structurally related to the complement C1 subcomponent, C1q, and seems to activate the complement system through an associated serine protease known as MASP (ref. 4) or p100 (ref. 5), which is similar to C1r and C1s of the classical pathway. MBL binds to specific carbohydrate structures found on the surface of a range of microorganisms, including bacteria, yeasts, parasitic protozoa and viruses⁶, and exhibits antibacterial activity through killing mediated by the terminal, lytic complement components⁷ or by promoting phagocytosis⁸. The level of MBL in plasma is genetically determined^{9–11}, and deficiency is associated with frequent infections in childhood^{12,13}, and possibly also in adults^{14,15} (for review, see ref. 6). We have now identified a new MBL-associated serine protease (MASP-2) which shows a striking homology with the previously reported MASP (MASP-1) and the two C1q-associated serine proteases C1r and C1s. Thus complement activation through MBL, like the classical pathway, involves two serine proteases and may antedate the development of the specific immune system of vertebrates.

Human plasma proteins and protein complexes that bind to carbohydrates in a calcium-dependent manner (lectins and lectin-associated proteins) were purified by affinity chromatography on mannan- and *N*-acetylglucosamine-derivatized Sepharose beads. This protein preparation was analysed by SDS-PAGE and blotting onto a PVDF-membrane. Development of the blot with chicken antibody raised against a bovine lectin preparation¹⁶ revealed a protein of relative molecular mass 52,000 (M_r 52K), in addition to MBL at 32K. The 52K band was subjected to amino-terminal amino-acid sequence analysis. The sequence showed similarity to

MAASP-2	TLPLGKWFPEVFGRLASGPPGPEYANDQERRWTLTAPPGYRLRLYFTHFDLHSLCEYDFVKLSSGAKVLATLGGQESTDTERAPGKDT	90
MAASP-1	HTVELNNMFGQIQSPGYPSDSSEVTWNITVPDGFRIKLYFMHFNLESSYLCEYDVKVEDQVLATFGRETTDETQTPGQEV	87
C1r	SIPIPQKLFGEVTSPLPKPKPNPFETTITVITVPTGYRVKLVFQDFLEPSEGCFDYVKISADKSLGRFGYQLGSPGNPFGKKE	87
C1s	EPTMYGELISPNYQAYPSEVKSWEDVPEGYIHLFYTHLDIELENGAYDSVQIISGDTTEGRGLGRSSNNPHSIVVEE	83
* * * * *		
MAASP-2	FYSLGSSLDITFRSDYSNEKP FTFGEAFYAEDIDEQ VAPGEA PTCDHHCHNLGGFYCSRAGYVLRHNRKTSALCS	170
MAASP-1	VLSPGSFMSTFRSDFSNEER FTFGDHYMAVDVDECK EREDEE LSCHDYCHNYIGGYCSRFYILHNDNRTRVECS	167
C1r	FMSQGNKMLLTFTDTSNEENGTMFYKGLAYQAVDLDEKASRSKSGEEDPQPCQHLCHNYVGGYFCSCRPYGLQEDRHSQAECS	177
C1s	FQVYNKLVQVIFKSDFSNEER FTGFAAYVATDINECT DFVD VPCHSFCHNFIGGYFCSCPPYFLHDDMKNGVNC	161
* * * * *		
MAASP-2	GQVFTQRSSSELSPFPRYPKLSSTYSISLEEGFVSILDFV ESFDVET HPETLCPYDFLKIQTDRREHGFPGKLTPLHR IETKS	256
MAASP-1	DNLFTQRTGVITSPDPFPNPKSSECLTYTIELEEGFMVNQFE DIFDIED HPEVFCPYDYIKIKVGPVLPFGFGEKAPET ISTQS	253
C1r	SRLYTRASGYISLEBYRSPYDPLRCHNYSIRVERGLTLHLKFL EPFDIDD HQQVHCPCYDQLQIYANGKNGEFCGKQRPED LDTSS	263
C1s	GDVFTALIGETASPNYPKYPENSCEFYQIRLEKGFQVVTLRREDFDVBAADAGNC LDSLVFVAGDRGPGYFGHGFPGLNIETKS	250
* * * * *		
MAASP-2	NTVTITFVDESQDHTGWKIHYSYTAQPCPYMAPPN GHVSPVQAKYILKDSFISFETGTYELLQGHLPKSFYAVQKQDGSWDRMPA	345
MAASP-1	HSVLILFHSNDSENROGRLSYRAAGNECELPQPPVH GKIEPSQAKYFPKQVLVSCTGYKVLKDNVEMDTFQIRCLKDGTSWNSKIPT	342
C1r	NAVDLFFTDRESGSRGWKRLRYTTIIEKCEPQKTLDEFTI IQNLQPCQYFRDYFIATCKQGYQLIEGNQVLHSTFAYQDDGTWHRAMPR	353
C1s	NALDIIFQDITGQKKGWKLRYHGDPMPEPKEDTNP SVWEPAKAKYVFRDVQITCLDGFVEVGRVGATSFYSTQCSQSNKSNKSLK	338
* * * * *		
MAASP-2	CSIVDCQPPDLSRVEYITGPGVITYKAVIQSCEETFTYM KVDNGKYVCEADGFWTSSSGEKSLPVECPVGLS ARTT	426
MAASP-1	CKIVDCQPPDLSRVEYITGPGVITYKAVIQSCEETFTYM NNNTGIYTCQAGVMMKVLGRSLPTCLPVGCLPKFSSKL	426
C1r	CKIKDCQPPRNLPGDFRFTYITMGVNTYKARIQYCEHYIKMQTRAGSRSESGVYVTCAGIWKNGQEKIIRCLPVGCKEPMVPEQ	443
C1s	QCPVDCQIPESINGKVB DPESTLFGSVIRYTCETPEYTYME NGGGGEYHCAGNSVWNEVLGPBLPKCPVPGVPREPFE	419
* * * * *		
MAASP-2	GGRIYGGQKAKGPFQVQLILGGTTA AGALLYDNWVLTAAH AVVEQKHDSALDIRMGLTKRLSPHYTQWSEAVFIHEG	507
MAASP-1	MARIFNGRPAQKGTTPWIAMLSHLNGQPCGSLGSSWIVTAACHLQSLDPKPTLRSDLLSPSD FKILGKHWRLSRDNELQHLG	515
C1r	RQRIGGQKAMGNFQVFTNTHGRG GGALLDGRWLTAAH TLYPKHEHAQSNAISLDVFLGHTNVEELMKLGNHP IRRV	523
C1s	KQRIGGSDADIKNFPQVFFDNPWA GGALINEXWVLTAAH VVEGNREFTMYVGSTSVQTSRLAKSKMLT PEHFVFIHG	498
* * * * *		
MAASP-2	YTHDAG FDMIDIALIKLNKVVINSNITPILPRKEAESFMRTDDIGTAGSWGLTQRGFLARNLMVVDIPVDHQCATAAYEK	589
MAASP-1	VKHTTLHPKYDPTFENDVALVELLESPLVNAFVMPICLP EGQQEGAMVIVSGWQKQFLQFPETIMIEIPIVDHSTQKQAY	599
C1r	SVHPDYRQDESYN FEQDIALLELENSVTLGNLLPILCLP DNDTFYDLGLMGVYSGFGVMEEK IAHLDHFRVLRLPVANPQACEN WLR	608
C1s	WKLLEV PEGRTN FDMIDIALVRLKDPVKGFTVSPILCPGTSSDYNLMGDLGLISGWGRTEKRDRAVRLKAARLPVAPLRKEKVKVE	586
* * * * *		
MAASP-2	PPYPRG SVTANMLCAGLESQGGKDCSRGDSGGALVFLDS ETERWFVGGIVSWGSMNCGEAGQYGVYTKVINYIPWIENIISDF	671
MAASP-1	APLKK KVTDRMICEKEGKQKDAESGDSGGPMVTLNR ERGQWLVGTVSWGD DQCKKDRYGVYSIHHNKDWIQRVTGVNR	680
C1r	GNRMD VFSCNNMFCAGHPSLKQDAQSGDSGGFAVRDP NTRDWATGVISWGI GCSRG YGFYTKVLNVYDWIKKEMEED	688
C1s	KPTADAEAYVFTPNMIGAGGK GMDSCKGDSSGGAFAVQPPNDKTKFYAAGLVSWGP QCGT YGLYTRVKNVVDWIMTKQENSTPRED	673
* * * * *		

Figure 2 Sequence alignment²¹ of the amino-acid sequences of MASP-2 (clone phl-4), MASP-1 (refs 17, 22), C1r (refs 23, 24) and C1s (refs 25, 26) beginning with the N-terminal amino acid of the mature proteins (the putative leader sequence, MRLTLGLLCCGSA, of MASP-2 is omitted). A bias of 6 was added to each term of the mutation data matrix (250PAMS) and a break penalty of 6 was used. Identical residues in all four species are indicated by asterisks. The beginning of the C1r/C1s-like domains, the EGF-like domain and the CCP domains are indicated above the sequences. Aligned cysteines are shaded. The potential cleavage site between Arg and Ile residues, which generates heavy and light chains, is identical to the site at which the serine protease domain starts. The three amino-acid residues that are essential for the active centre in serine proteases (His 468, Asp 517 and Ser 618), are indicated by diamonds. The cysteines in the histidine loop of MASP-1 are marked (inverted triangles). Sequences obtained from peptides are underlined.

that of the previously described MASP (MASP-1). Antibody raised against a synthetic peptide representing the 19 N-terminal amino acids (anti-N'MASP-2 antiserum) recognized the 52K molecule as well as a molecule with a mobility corresponding to 20K (Fig. 1, lane 1). Under non-reducing conditions, a polypeptide of 76K was detected with anti-N'MASP-2 antiserum (Fig. 1, lane 2), indicating the presence of interchain disulphide bonds. The 20K polypeptide had the same N-terminal sequence as the 52K polypeptide and is likely to represent a truncated form of it. The directly determined amino-acid sequences (N-terminal, as well as those of internal peptides) are indicated in Fig. 2. Two-dimensional SDS-PAGE, in which the first dimension was under non-reducing conditions and the second dimension under reducing conditions, showed the 52K polypeptide to be part of a disulphide-linked protein of M_r 76K. A polypeptide of 31K (Fig. 1, lane 3), probably representing the remainder of the protein, was also recognized as part of the 76K protein by an antiserum (anti-C'MASP-2) raised against synthetic peptides from the C-terminal sequence of the protein (determined by cDNA sequencing; see below). The 76K band seen with the anti-N'MASP-2 antibody under non-reducing conditions was also recognized by the anti-C'MASP-2 antibody (Fig. 1, lane 4).

The liver is the primary site of synthesis of C1r, C1s and MASP-1. Thus, RNA from liver was used as the template for polymerase chain reaction with reverse transcription (RT-PCR), with primers deduced from the peptide sequences. The nucleotide sequence of the resulting 300-base-pair (bp) RT-PCR product contained an open reading frame (ORF), with a deduced amino-acid sequence confirming the sequences of the peptides from which the primers were derived, as well as that of another of the sequenced peptides. The RT-PCR-derived product was used as a probe for screening a human liver cDNA library, revealing 16 hybridizing clones, of which the four longest (phl-1, 2, 3 and 4) were sequenced in full. Sequence

analysis revealed that all four clones represent reverse transcripts of the same novel human mRNA species. The longest clone, phl-4, comprises 2,475 bp starting with a 5' untranslated region of 36 bp, followed by an ORF of 2,061 bp and a 3' untranslated region of 378 bp, ending with a poly(A) tail. The nucleotide sequence of phl-4 is deposited at the EMBL nucleotide sequence data base (accession no. Y09926). Although the sequence of phl-1 and -2 were in total agreement with phl-4, the nucleotide sequence of phl-3 differs from phl-4 at two positions, a transversion at nucleotide position 1,147 (G to T) and a transition at position 1,515 (C to T). The first change leads to the replacement of Asp 356 with tyrosine. All clones were isolated from a liver library transcribed from RNA from a single donor, and the observed difference may represent polymorphism in the MASP-2 gene, or is due to an error created during construction of the library.

The amino-acid sequences of the N terminus as well as all sequenced peptides were identified in the sequence deduced from clone phl-4. The ORF encodes a polypeptide chain of 686 amino acids, including a signal peptide of 15 residues (Fig. 2). Omitting the signal peptide, the calculated M_r is 74,153, in agreement with the 76K observed on SDS-PAGE (Fig. 1), the isoelectric point is 5.43, and the molar extinction coefficient is 113,640 (absorbance at 280 nm of 1.54 at 1 mg ml⁻¹). In contrast to MASP-1, the sequence contains no sites for N-linked glycosylation. The three amino-acid residues that are essential for the active centre in serine proteases (His 468, Asp 517 and Ser 618) are present. The amino-acid sequence is homologous to that of MASP-1, C1r and C1s (Fig. 2). These four proteins have in common their domain organization, featuring one C1r/C1s-like domain, one epidermal-growth-factor-like (EGF-like) domain, followed by a second C1r/C1s-like domain, two complement control protein (CCP) domains, and a serine protease domain. The key residues involved in the calcium-binding

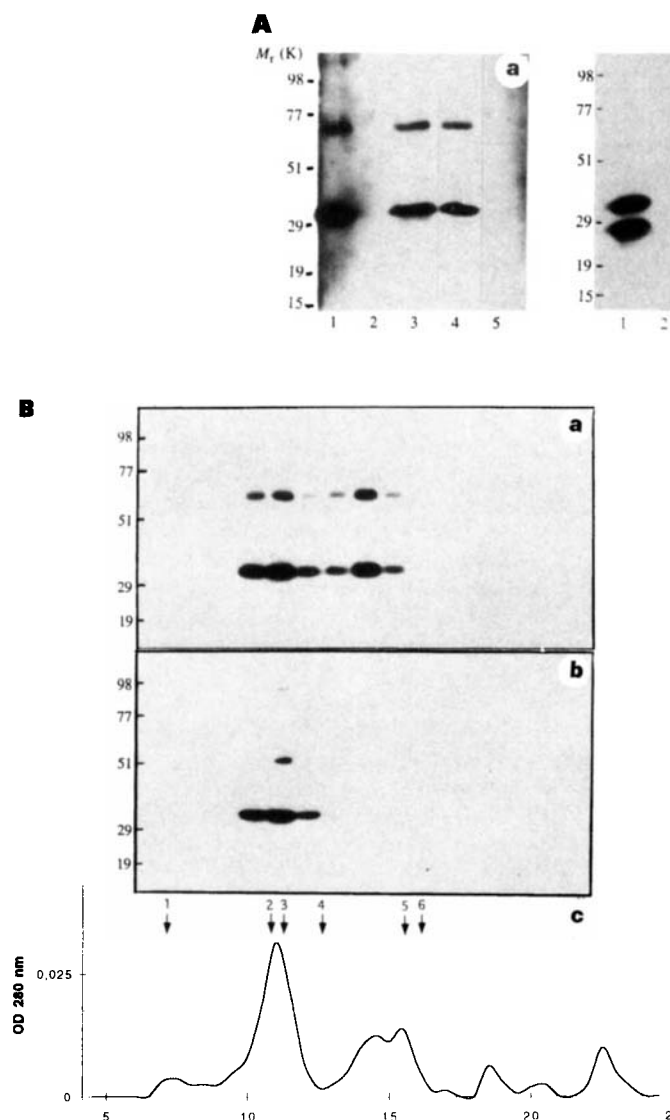


Figure 3 Molecular complex between MBL, MASP-1 and MASP-2. **A**, Affinity purification of MBL/MASP complexes. Microtitre wells were coated with monoclonal anti-MBL or control monoclonal murine IgG1, incubated with either one of two different lectin preparations (1 and 2), and the bound proteins were eluted and analysed by SDS-PAGE under reducing conditions and by western blotting. Blot **a** was developed with anti-MBL antibody, blot **b** with anti-C' MASP-1 antibody, and blot **c** with anti-N' MASP-2 antibody. Lane 1, unfractionated lectin, preparation (1); lanes 3 and 4, eluates from wells coated with anti-MBL antibody and incubated with lectin preparations (2) and (1), respectively; lanes 2 and 5, eluates from wells coated with normal IgG and incubated with lectin preparations (2) and (1), respectively. **B**, The lectin preparation was subjected to gel-permeation chromatography on a Superose-6 column in buffer containing calcium. **a**, Analysis of the fractions by western blotting using monoclonal anti-MBL antibody. The band at ~60K is seen in all MBL preparations and is recognized by all anti-MBL antibodies tested (monoclonal as well as polyclonal) and probably represents a non-reducible dimer of the 32K polypeptide chain. **b**, Same analysis, using anti-N' MASP-2 antibody (developing the upper band of 52K followed by anti-C' MASP-1 antibody (developing the lower band of 31K). For technical reasons, the 20K truncated MASP-2 is not seen here as the blot was partially stripped between incubations with anti-MASP-2 and anti-MASP-1. **c**, Numbered arrows on the chromatography elution profile indicate the void volume (1) and the elution positions for the marker proteins IgM (2), Clq (3), thyroglobulin (4), IgG (5) and serum albumin (6).

motif in the EGF-like domains are present in the sequence, as well as in MASP-1, C1r and C1s. In addition, the substrate specificity-related residue, six residues before the active-site serine, is aspartic acid in all four proteins. MASP-1, C1r and C1s are all activated by cleavage of the peptide bond between the residues arginine and isoleucine located between the second CCP domain and the serine protease domain. The resulting polypeptide chains (the larger is referred to as the 'heavy' chain and the smaller as 'light' chain) are held together by a disulphide bond. By analogy, our results indicate that the 52K polypeptide, recognized by antibody against the N-terminal of MASP-2 after SDS-PAGE under reducing conditions, is the heavy chain of MASP-2, whereas the 31K polypeptide, recognized by antibody against the C terminus of MASP-2, is the light chain. As seen in Fig. 2, arginine and isoleucine are present in MASP-2 at the expected positions between the second CCP domain and the protease domain.

Identities and similarities between the four proteins were studied from the sequence alignment shown in Fig. 2. Identity between the four proteins, all in the range of 39 to 45%, gives no clue as to functional relatedness. The similarity scores (taking into account residues of a similar nature as well as identical residues) between the proteins are between 39 and 52%, with the least similarity being between MASP-1 and C1s (39%) and the highest similarity between MASP-1 and C1r (52%) and between MASP-1 and MASP-2 (52%). MASP-2 shows comparable similarity with C1r (46%) and C1s

(47%). The similarity score between C1s and MASP-2 is significantly higher than that between C1s and MASP-1, whereas MASP-1 is more similar to C1r than to C1s, suggesting that MASP-2, like C1s, could be the C2 and C4 cleaving enzyme. Several features of the sequences suggest that MASP-2, C1r and C1s have evolved by gene duplication and divergence from a MASP-1 ancestor. Only the MASP-1 sequence contains the histidine loop characteristic of trypsin-like serine proteases¹⁷. The active-site serine is encoded by a TCN codon (where N is A, T, G or C) in MASP-1 as in most serine proteases, whereas in MASP-2, C1r and C1s it is encoded by an AGY codon (where Y is T or C). In most serine proteases, including MASP-1, a proline residue is found at the third position downstream from the active-site serine, whereas a different amino acid is found in MASP-2, C1s and C1r (alanine in MASP-2 and C1s, valine in C1r). From these analogies, the catalytic domain of MASP-2 may be predicted to be encoded by a single exon as in C1r and C1s, whereas most other serine proteases, including MASP-1¹⁸, have split exons.

To investigate the association between MBL, MASP-1 and MASP-2, the lectin preparation was incubated in microtitre wells coated with monoclonal anti-MBL antibody, or, as a negative control, wells coated with nonspecific monoclonal immunoglobulin of the same subclass. The proteins captured by the antibody were eluted and analysed by SDS-PAGE and western blotting. The results (Fig. 3A) show that the anti-MBL antibody, in addition to binding MBL,

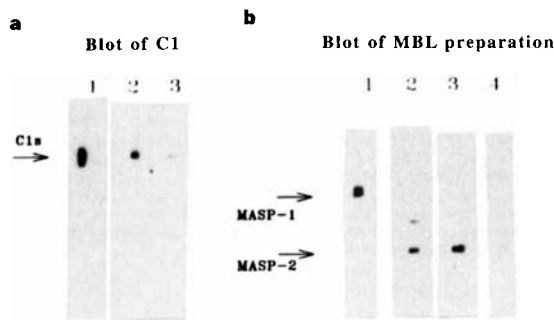


Figure 4 Activation of C4 by C1s and MASP-2 on western blots. **a**, Western blot of C1 separated under non-reducing conditions, without heating the sample before electrophoresis. Lane 1 was developed with anti-C1s antibody; lane 2 was incubated with human serum followed by anti-C4 antibody; lane 3 was as lane 2, except for the presence of serine-protease inhibitors during the incubation with serum. **b**, Western blot of an MBL preparation separated as in **a**. Lane 1 was developed with anti-N¹MAASP-1; lane 2 with anti-N¹MAASP-2; lane 3 was incubated with human serum at 37°C followed by anti-C4; in lane 4, the blot was preincubated with serine-protease inhibitors and the incubation with serum was also in the presence of inhibitors. MASP-1 shows a higher *M_r* than MASP-2 owing to glycosylation¹⁷ and a polypeptide chain that is 9 amino acids longer.

captures both MASP-1 and MASP-2. Fractions from gel-permeation chromatography of the lectin preparation on Superose-6B CL were analysed for MBL, MASP-1 and MASP-2 (Fig. 3B). MBL was eluted in a main peak at a volume (*V_e*) corresponding to an *M_r* of 750K, with a smaller peak at 350K. MASP-1 and MASP-2 were found by western blotting to co-elute largely with the high-molecular-mass MBL. When the MBL preparation was chromatographed at pH 5, no MASP-1 or MASP-2 was associated with the MBL peak, as already reported for MASP-1 (ref. 19).

The classical complement activation pathway, like the MBL-initiated pathway, involves the generation of a C3-converting complex, C4b2b, through enzymatic activation of C4 and C2. In the C1 complex (C1q₂s₂), this specific protease activity is exhibited by C1s after activation of this enzyme by C1r. When C4 is activated, its reactive thiol ester is exposed and C4b binds covalently to nearby amino or hydroxyl groups. The C4-activating abilities of MASP-1 and MASP-2 were compared. This was accomplished by separating an MBL/MAASP preparation (prepared in the absence of enzyme inhibitors) by SDS-PAGE under non-reducing conditions, followed by blotting onto a PVDF membrane in the absence of SDS. The blot was examined for C4 converting activity by incubation with human serum at 37°C, followed by detection of deposited C4b with anti-C4 antibodies (Fig. 4). C4 was deposited at a position corresponding to the MASP-2 band, whereas no C4 deposition occurred at positions corresponding to MASP-1. MASP-1 was present in the activated state, as shown by SDS-PAGE under reducing conditions, where it appears as two bands at ~30K and 70K, respectively (results not shown). This C4 activation and deposition was inhibited by enzyme inhibitors (Fig. 4). No C4-stimulating activity could be detected when MBL/MAASP was prepared in the presence of enzyme inhibitors added throughout the purification procedure. A preparation of C1 was similarly analysed, and C4 deposition, which could be inhibited by enzyme inhibitors, was observed at a position corresponding to C1r and C1s, which are not separated under our conditions.

Our results emphasize the similarity between complement activation through the MBL, or 'MBLectin' pathway of the innate immune system and the classical pathway of complement activation (Fig. 5). In both cases, the initiating molecular complexes are composed of an oligomeric ligand-binding component (MBL or

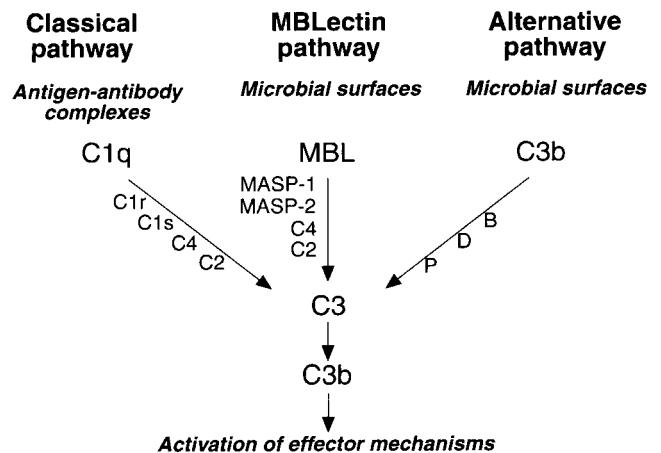


Figure 5 The three pathways of complement activation. Activation through the classical pathway is associated with the specific immune response found only in vertebrates, whereas the MBL, or 'MBLectin', pathway and the alternative pathway rely on innate recognition of foreign organisms and so probably predate the evolution of the specific immune system. All pathways converge on the activation of the central component C3 into C3b, which binds covalently to the microbial surface and mediates the effector functions of complement.

C1q, respectively) which, on reacting with ligands, activates the two associated serine proteases (MASP-1 and MASP-2 or C1r and C1s, respectively). Elucidation of the precise composition and stoichiometry of the MBL/MASP-1/MASP-2 complex, and the enzymatic properties of the two proteases, requires further investigation. □

Methods

Purification of plasma lectins and associated molecules. Pooled CPD plasma (2.5 l), diluted with buffer containing EDTA and enzyme inhibitors, was passed through Sepharose-2B CL and mannan-Sepharose. A thrombin inhibitor, PPACK (D-phenylalanyl-prolyl-arginyl-chloromethyl ketone), and CaCl₂ were added. The pool was passed through Sepharose-2B CL and mannan-Sepharose, and the proteins binding calcium-dependently to mannan-Sepharose were eluted with EDTA-containing buffer. The eluate was recalcified, passed through a GlcNAc-Sepharose column, which was eluted to yield 20 ml of lectin preparation.

Antibodies. Animals, primed with BCG (bacillus Calmette-Guérin vaccine), were immunized with synthetic peptides coupled to PPD (tuberculin-purified protein derivative) according to C. Koch (State Serum Institute, Copenhagen). Antibodies anti-N¹MAASP-1, anti-C¹MAASP-1 and anti-N¹MAASP-2 were from rabbits immunized with peptides corresponding to the first 19 amino-acid residues of MASP-1, the last 19 amino-acid residues of MASP-1, and the first 19 amino-acid residues of MASP-2, respectively. Chicken anti-C¹MAASP-2 antibody was from chickens immunized with a mixture of two peptides representing sequences in the C-terminal part of MASP-2 (residues 505–523 and 538–556). All peptides had an additional C-terminal cysteine for coupling. Antibody and normal chicken IgG was purified from yolk²⁰. Monoclonal anti-MBL antibody, IgG1-κ (clone 131-1) and control IgG1-κ (clone MOPC21) were purified by protein-A affinity chromatography. F(ab')₂ rabbit anti-human C4 and F(ab')₂ rabbit anti-human C1q were produced by pepsin digestion of rabbit anti-human C4 and rabbit anti-human C1q (Dako, Glostrup, Denmark). For staining of western blots, antibodies were used at 1 μg ml⁻¹. Bound chicken antibody was visualized with rabbit anti-chicken IgG followed by peroxidase-labelled goat anti-rabbit IgG and developed by enhanced chemiluminescence. Bound mouse and rabbit antibodies were visualized with peroxidase-labelled rabbit anti-mouse IgG and peroxidase-labelled goat anti-rabbit IgG, respectively.

Amino-acid sequencing of the 52K and the 20K polypeptides. The lectin preparation was concentrated, run on SDS-PAGE and transferred to a PVDF membrane. Two strips were developed with anti-bovine lectin antibody¹⁶. The

rest of the blot was stained with Coomassie brilliant blue. The band corresponding to the immunostained 52K band was cut out and sequenced on an Applied Biosystems protein sequencer. After production of anti-N'MASP-2 antibody, this was used on a similar western blot and the N termini of the proteins in the 52K and the 20K bands, developed with this antibody, were sequenced. Peptides were prepared by trypsin digestion of the proteins in the two bands from another blot, fractionated by reverse-phase chromatography and the peptides in the principal peaks were sequenced.

C4 activation on western blots. C1 was purified by incubating IgG-coupled Sepharose beads with human serum at 4°C. The beads were washed and incubated at 37°C for 30 min for activation of C1r and C1s. The beads were suspended in non-reducing sample buffer and, without boiling, run on SDS-PAGE, then blotted in the absence of SDS. A similar blot was made of an MBL preparation produced in the absence of enzyme inhibitors. Strips of the blots were incubated for 30 min at 37°C with 1.1% (v/v) human MBL-deficient serum, depleted of C1q by fractionation on Biorex-70. Blots were developed with biotinylated F(ab')₂ anti-C4 antibody, followed by peroxidase-labelled streptavidin and luminescence reagent. Parallel blots were treated with a serine-protease inhibitor (aminoethylbenzenesulphonyl fluoride), which was also present during incubation with serum. Other strips were directly developed with antibodies.

Cloning of the MASP-2 cDNA. First-strand synthesis of cDNA was carried out with 1.3 µg human liver RNA using a First-Strand cDNA synthesis kit (Pharmacia). PCR was done on this using degenerated sense and antisense primer derived from the amino-acid sequences EYANDQER and KPFTGFEA, respectively. The PCR program consisted of 1 cycle of annealing at 50°C; 1 cycle of annealing at 55°C, and 33 cycles of annealing at 60°C. The resulting 300-bp PCR product was cloned into the *E. coli* plasmid pCRII using the TA-cloning kit (Invitrogen) and the nucleotide sequence of the insert was determined. The insert of this plasmid was used as the radioactively labelled probe for screening a total of 8×10^5 clones in a commercial human liver library (Stratagene).

Received 30 September 1996; accepted 4 February 1997.

- Law, S. K. A. & Reid, K. B. M. *Complement* 2nd edn (ed. Male, D.) 1–88 (IRL, Oxford, 1996).
- Ikeda, K., Sannoh, T., Kawasaki, N., Kawasaki, T. & Yamashina, I. Serum lectin with known structure activates complement through the classical pathway. *J. Biol. Chem.* **262**, 7451–7454 (1987).
- Kawasaki, T., Etoh, R. & Yamashina, I. Isolation and characterisation of a mannan-binding protein from rabbit liver. *Biochem. Biophys. Res. Commun.* **81**, 1018–1204 (1978).
- Matsushita, M. & Fujita, T. Activation of the classical complement pathway by mannose-binding protein in association with a novel C1s-like serine protease. *J. Exp. Med.* **176**, 1497–1502 (1992).
- Ji, Y.-H. *et al.* Activation of the C4 and C2 components of complement by a proteinase in serum bactericidal factor, Ra reactive factor. *J. Immunol.* **150**, 571–578 (1993).
- Turner, M. W. Mannose-binding lectin: the pluripotent molecule of the innate immune system. *Immunol. Today* **17**, 532–540 (1996).
- Kawasaki, N., Kawasaki, T. & Yamashina, I. A Serum lectin (mannan-binding protein) has complement-dependent bactericidal activity. *J. Biochem.* **106**, 483–489 (1989).
- Kuhlman, M., Joiner, K. & Ezekowitz, R. A. B. The human mannose-binding protein functions as an opsonin. *J. Exp. Med.* **169**, 1733–1745 (1989).
- Sumiya, M. *et al.* Molecular basis of opsonic defect in immunodeficient children. *Lancet* **337**, 1569–1570 (1991).
- Lipscombe, R. J. *et al.* High frequencies in African and non-African populations of independent mutations in the mannose binding protein gene. *Hum. Mol. Genet.* **1**, 709–715 (1992).
- Madsen, H. O. *et al.* A new frequent allele is the missing link in the structural polymorphism of the human mannan-binding protein. *Immunogenetics* **40**, 37–44 (1994).
- Super, M., Thiel, S., Lu, J., Levinsky, R. J. & Turner, M. W. Association of low levels of mannan-binding protein with a common defect of opsonisation. *Lancet* **ii**, 1236–1239 (1989).
- Garred, P., Madsen, H. O., Hofmann, B. & Sveigaard, A. Increased frequency of homozygosity of abnormal mannan-binding protein alleles in patients with suspected immunodeficiency. *Lancet* **346**, 941–943 (1995).
- Summerfield, J. A. *et al.* Mannose binding protein gene mutations associated with unusual and severe infections in adults. *Lancet* **345**, 886–889 (1995).
- Nielsen, S. L., Andersen, P. L., Koch, C., Jensenius, J. C. & Thiel, S. The level of the serum opsonin, mannan-binding protein in HIV-1 antibody-positive patients. *Clin. Exp. Immunol.* **100**, 219–222 (1995).
- Baattrup, G., Thiel, S., Isager, H., Svevag, S. E. & Jensenius, J. C. Demonstration in human plasma of a lectin activity analogous to that of bovine conglutinin. *Scand. J. Immunol.* **26**, 355–361 (1987).
- Sato, T., Endo, Y., Matsushita, M. & Fujita, T. Molecular characterization of a novel serine protease involved in activation of the complement system by mannose-binding protein. *Int. Immunol.* **6**, 665–669 (1994).
- Endo, Y., Sato, T., Matsushita, M. & Fujita, T. Exon structure of the gene encoding the human mannose-binding protein-associated serine protease light chain: comparison with complement C1r and C1s genes. *Int. Immunol.* **9**, 1355–1358 (1996).
- Tan, S. M. *et al.* Improvements on the purification of mannan-binding lectin and demonstration of its Ca²⁺-independent association with a C1s-like serine protease. *Biochem. J.* **319**, 329–332 (1996).
- Jensenius, J. C., Andersen, I., Hau, J., Crone, M. & Koch, C. Eggs: conveniently packaged antibodies. Methods for purification of yolk IgG. *J. Immunol. Meth.* **46**, 63–66 (1981).
- Barton, G. J. Protein multiple sequence alignment and flexible pattern matching. *Meth. Enzymol.* **183**, 403–428 (1990).
- Takada, F., Takayama, Y., Hatsuse, H. & Kawakami, M. A new member of the C1s family of complement proteins found in a bactericidal factor, Ra-reactive factor, in human serum. *Biochem. Biophys. Res. Commun.* **196**, 1003–1009 (1993).

- Journat, A. & Tosi, M. Cloning and sequencing of full-length cDNA encoding the precursor of human complement component C1r. *Biochem. J.* **240**, 783–787 (1986).
- Lytus, S. P., Kurachi, K., Sakariassen, K. S. & Davie, E. W. Nucleotide sequence of cDNA coding for human complement C1r. *Biochemistry* **25**, 4855–4863 (1986).
- Mackinnon, C. M., Carter, P. E., Smyth, S. J., Dunbar, B. & Fothergill, J. E. Molecular cloning of cDNA for human complement component C1s. The complete amino acid sequence. *Eur. J. Biochem.* **169**, 547–553 (1987).
- Tosi, M., Duponchel, C., Meo, T. & Julier, C. Complete cDNA sequence of human complement C1s and close physical linkage of the homologous genes C1s and C1r. *Biochemistry* **26**, 8516–8524 (1987).

Acknowledgements. We thank A. Day for alignment examination, and A. G. Hansen, J. Lovmand and J. Broholm for technical assistance. The term 'MBlectin pathway' was suggested by C. Janeway. This investigation was supported by grants from the Danish MRC, the Danish National Research Foundation, the Novo Nordisk Foundation, the Alfred Benzon Foundation, the Plasmid Foundation, FEBS, the Wellcome Trust and the Deutsche Forschungsgemeinschaft.

Correspondence and requests for materials should be addressed to S.T. (e-mail: jensenius@svfcd.aau.dk).

Inhibition of natural killer cells by a cytomegalovirus MHC class I homologue *in vivo*

H. E. Farrell, H. Vally, D. M. Lynch, P. Fleming, G. R. Shellam, A. A. Scalzo & N. J. Davis-Poynter

Department of Microbiology, University of Western Australia, Queen Elizabeth II Medical Centre, Nedlands 6907, Australia

Herpesviruses, such as murine and human cytomegalovirus (MCMV and HCMV), can establish a persistent infection within the host and have diverse mechanisms as protection from host immune defences¹. Several herpesvirus genes that are homologous to host immune modulators have been identified, and are implicated in viral evasion of the host immune response^{2,3}. The discovery of a viral major histocompatibility complex (MHC) class I homologue, encoded by HCMV⁴, led to speculation that it might function as an immune modulator and disrupt presentation of peptides by MHC class I to cytotoxic T cells⁵. However, there is no evidence concerning the biological significance of this gene during viral infection. Recent analysis of the MCMV genome has also demonstrated the presence of a MHC class I homologue⁶. Here we show that a recombinant MCMV, in which the gene encoding the class I homologue has been disrupted, has severely restricted replication during the acute stage of infection compared with wild-type MCMV. We demonstrate by *in vivo* depletion studies that natural killer (NK) cells are responsible for the attenuated phenotype of the mutant. Thus the viral MHC class I homologue contributes to immune evasion through interference with NK cell-mediated clearance.

The sequence of the MCMV MHC class I homologue (designated m144) is remarkably similar to that of MHC class I proteins (Fig. 1). It is predicted to encode a 383 amino-acid type 1 membrane glycoprotein that bears sequence homology with the α1, α2 and α3 domains of MHC class I heavy chains^{7,8}. The highest level of sequence identity between m144 and MHC class I proteins is observed in the putative α3 domain (32% identity), including the cysteines known to form internal disulphide bonds in cellular class I proteins. Least conserved is the predicted α2 domain of m144 (20% identity), which is truncated compared to MHC class I proteins and the HCMV MHC class I homologue encoded by the UL18 gene. Low homology is also observed in the cytoplasmic region of m144, which is considerably longer than that of MHC class I proteins. Unlike UL18, which has 13 predicted N-linked glycosylation sites, m144 has only 4, which more closely approximates the number (1–3) found in HLA and H-2 class I molecules. UL18 and m144 are less conserved with each other than with MHC class I molecules from their respective species, suggesting that the genes were acquired independently or have diverged from an ancestral counterpart during adaptation of the viruses to their natural host.



Figure 1 Alignment of the predicted amino-acid sequence of m144 (strain K181) with the α -chain of murine MHC class I (muMHC), HCMV UL18, and the α -chain of human MHC class I (huMHC). For clarity, only one representative murine and human MHC class I sequence is shown. Alignments are grouped into putative signal, $\alpha 1$, $\alpha 2$, $\alpha 3$, transmembrane (Tm) and cytoplasmic (cyto) domains. Identical amino acids are marked with asterisks, and similar amino acids are marked with dots. Residues conserved between m144 and UL18 alone are shown on the line

separating the two viral sequences; the conservation between the viral and cellular class I proteins is shown above (for m144 and muMHC) or below (for UL18 and huMHC) each of the paired alignments. The similarity with MHC class I proteins was determined using FASTA²⁹, and sequences were aligned using CLUSTALW³⁰. Potential amino-linked glycosylation sites (N-X-T/S) are underlined; conserved cysteine residues are shown in bold type.

We examined the role of the m144 gene product *in vivo* by comparing the ability of a MCMV mutant with a disrupted m144 open reading frame (ORF) to replicate with that of wild-type MCMV. The m144 ORF was disrupted by the insertion of a dual *gpt/lacZ* cassette (Fig. 2a), which was confirmed by Southern blotting (Fig. 2b). Similar to results described previously for a HCMV UL18 deletion mutant⁹, the m144 deletion mutant (designated K Δ 144) replicated to wild-type levels in fibroblasts *in vitro* (Fig. 3a). *In vivo*, however, K Δ 144 replicated poorly in the visceral organs of susceptible BALB/c mice during the acute stage post-infection (days 2–6), with titres up to 400-fold lower than those of wild-type-infected mice (Fig. 3b). A MCMV recombinant for which the lesion in m144 was repaired with wild-type MCMV sequence (that is, a m144 'revertant', designated K Δ 144r) replicated to wild-type levels *in vivo*, confirming that the phenotype observed for K Δ 144 was caused by disruption of m144 rather than adventitious mutations elsewhere in the MCMV genome. Levels of replication in the salivary gland, which is a site of persistence for both MCMV and HCMV in their natural hosts, was unaffected by the disruption of m144.

We investigated whether the restricted replication of K Δ 144 was due to an altered ability of the host to clear the virus during the first few days post-infection. The phenotype of K Δ 144 is consistent with the timing of a NK cell response. NK cells are induced shortly after MCMV infection, although the contribution of NK cells towards MCMV clearance varies according to the genetic constitution of the host^{10–12}. Indeed, the significance of interactions between the NK cell and the host cell in the early replication of MCMV has been

shown by the identification of an autosomal dominant host gene, *Cmv1*, which contributes to resistance to MCMV infection. This locus mediates its effect through the action of NK cells and maps closely to the NK gene complex^{13,14}. To assess the influence of NK cells in limiting the replication of K Δ 144, we compared virus titres from mice depleted of cells expressing asialo GM₁ and control BALB/c mice infected with 10⁴ plaque-forming units of either wild-type MCMV (strain K181) or K Δ 144. Both K181 and K Δ 144 induce strong cytolytic NK cell responses after infection (Fig. 4a), which are ablated by *in vivo* pretreatment with anti-asialo GM₁ antiserum. Although the presence or absence of NK cells in K181-infected mice had no effect upon virus titres in the spleen at this dose, and had only a moderate effect on replication in the liver, K Δ 144 replication in both organs was substantially increased in NK cell-depleted animals (Fig. 4b). Expression of asialo GM₁ is not restricted to NK cells, and it is possible that the effects observed upon K Δ 144 replication following treatment with anti-asialo GM₁ may be due to depletion of cell types other than NK cells, including T cells. However, depletion of both CD4⁺ and CD8⁺ T cells was found to have no significant effect on the replication of K Δ 144 compared with non-depleted animals. These results demonstrate that, although equivalent levels of NK cell cytolytic activity are induced by K181 and K Δ 144, the *in vivo* replication of the m144 deletion mutant was more efficiently restricted in a NK cell-dependent manner.

The most specific serological marker of mouse NK cells is the NK1.1 antigen, which is recognized by the monoclonal antibody PK136 (ref. 15). The NK1.1 antigen is not expressed in BALB/c

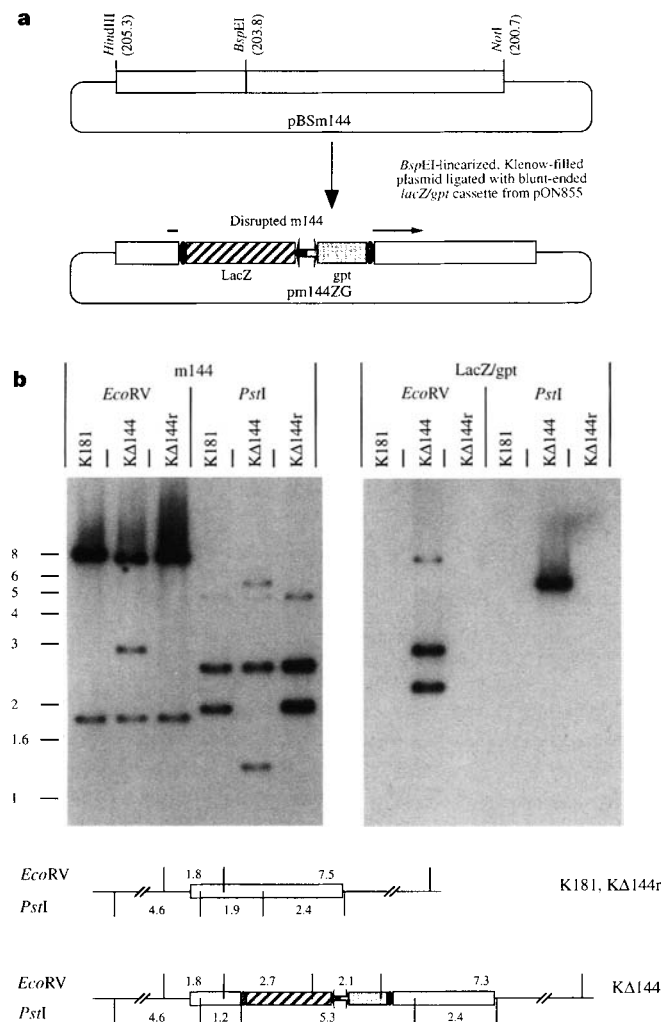


Figure 2 Production of recombinant MCMV deleted of m144 expression. **a**, Plasmids used. **b**, Southern blot analysis of wild-type MCMV (K181), KΔ144 and KΔ144r. DNA from MEF infected with either K181, KΔ144 or KΔ144r was digested with *EcoRV* or *PstI* and probed with either the 4.6-kb *HindIII*/*NorI* fragment of pBSm144 (left) or the *lacZ*/*gpt* cassette of pON855 (right). The predicted fragment lengths for the various viruses are displayed beneath the autoradiograms. The sizes (kb) and positions of DNA size markers are shown to the left of the figure.

mice, and so we performed *in vivo* depletion studies in the congenic mouse strain BALB.B6-*Cmv1*⁺ (ref. 14). This strain possesses the BALB/c genotype, except for a region of chromosome 6 that is derived from the C57BL/6 mouse strain and encompasses the NK gene complex (including the *mNKR-P1c* gene that encodes the NK1.1 antigen) and the *Cmv1* resistance gene; the latter confers low virus titres in the spleens of MCMV-infected mice through the action of NK cells¹³. In this strain, replication of KΔm144 was not detected in either the spleen or the liver during the first 6 days after MCMV infection (data not shown). Consistent with the previous results following anti-asialo GM₁ treatment of BALB/c mice, depletion of NK1.1-positive cells resulted in a significant increase in replication of KΔm144 in the spleen and liver (Fig. 4c). Taken together, our data indicate that NK cells are more effective at restricting replication of KΔm144 than of wild-type MCMV.

Some of the MHC class I-specific receptors expressed by NK cells

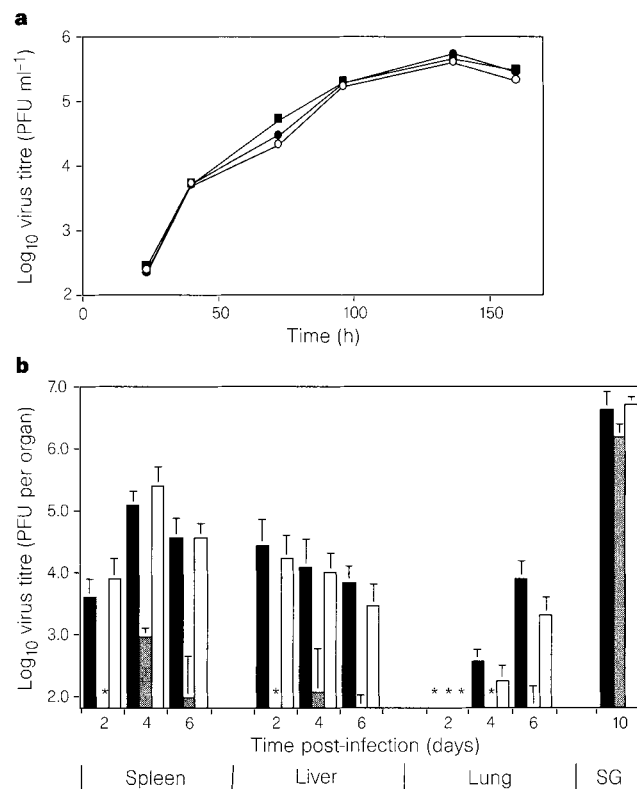


Figure 3 Analysis of the phenotype of MCMV m144. **a**, *In vitro* growth curves of wild-type MCMV strain K181 (filled squares), KΔ144 (filled circles) and KΔ144r (open circles) in primary mouse embryo fibroblast (MEF) cultures. Each point represents the mean of three independent infections, and is expressed as log₁₀ plaque forming units (PFU) ml⁻¹. Standard deviations were less than 15% of the mean values, and are not shown. **b**, Growth of wild-type (black), KΔ144 (hatched) and KΔ144r (white bars) *in vivo* after an intraperitoneal challenge of 10⁴ PFU. Each group represents the log₁₀ of the mean PFU per organ (±s.d.) from four mice. Groups in which no virus was detected in any of the mice tested are indicated by asterisks. The organs collected are indicated beneath the graph: SG denotes salivary gland.

have been identified^{16–18}. In contrast to CD8⁺ T cells, NK cells are inhibited, rather than activated, by target cell expression of MHC class I (ref. 19). Expression of low levels of MHC class I is therefore often associated with increased susceptibility to NK cell-mediated lysis. Because both MCMV and HCMV cause downregulation of MHC class I expression on the cell surface^{20–22}, this would potentially make infected cells susceptible targets for NK cell-mediated immunity. We propose that m144 mimics cellular MHC class I proteins and acts to inhibit the host NK cell response, thus preventing the early clearance of infected cells. Similarly, HCMV UL18 has been postulated as a possible inhibitor of NK cell activation as it mimics the properties of MHC class I molecules with regard to binding of β2 microglobulin²³ and endogenous peptide²⁴. However, analysis of the possible function of UL18 as an immune modulator *in vivo* is not possible owing to the lack of a suitable animal model system. The use of MCMV recombinants

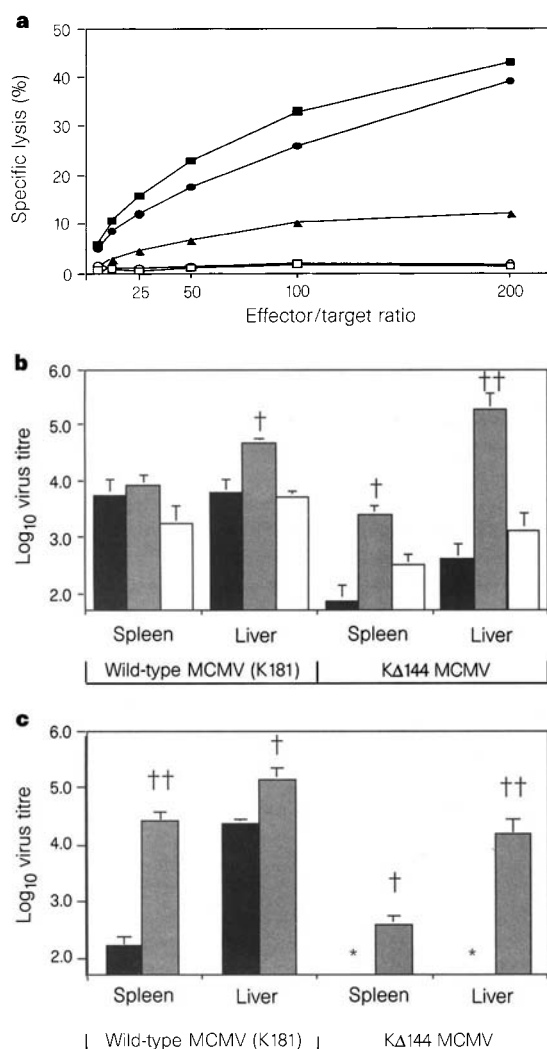


Figure 4 Clearance of MCMV in the absence of the MHC class I homologue is mediated by NK cells. **a**, NK cytotoxicity of spleen cells from MCMV-infected and uninfected mice against YAC-1 targets. Increased levels of NK cell activity were observed in BALB/c mice after infection with either wild-type K181 (filled squares) or KΔ144 (filled circles) compared with uninfected controls (filled triangles). NK cell cytotoxicity is fully ablated by *in vivo* pretreatment with anti-asialoGM₁ in both wild-type (open squares) and KΔ144-infected (open circles) mice. Each point represents the average cytotoxicity of 3 mice; s.d. were less than 15% of the mean values and are not shown. **b**, Quantification of MCMV titres in wild-type K181- and KΔ144-infected BALB/c mice pretreated *in vivo* with antibodies to either asialo GM₁ (hatched), CD4 and CD8 (white) or saline (black). Titres were determined at 3 days post-infection. Each group represents the log₁₀ of the mean PFU per organ ($\pm$ s.d.) from four mice. The *in vivo* phenotype of KΔ144 is ablated in NK cell-depleted, but not in CD4⁺/CD8⁺ T-cell-depleted, mice. Where differences in virus titre between saline-treated or lymphocyte-depleted animals are statistically significant, this is indicated: †*P* < 0.01; ††*P* < 0.001. **c**, Quantification of MCMV titres in wild-type K181- and KΔ144-infected BALB.B6-Cmv1^{-/-} mice pretreated *in vivo* with monoclonal antibody PK136 (anti-NK1.1; hatched) or saline (black). Titres were determined at 3 days post-infection. Each group represents the log₁₀ of the mean PFU per organ ($\pm$ s.d.) from four mice. Groups in which no virus was detected in any of the mice tested are indicated by asterisks. Depletion of NK1.1⁺ cells resulted in statistically significant increases in virus titres compared with saline-treated animals, as indicated: †*P* < 0.01; ††*P* < 0.001.

expressing mutated forms of m144 will enable a detailed *in vivo* characterization of the evasion of NK cell-mediated immunity. □

Methods

Plasmid construction. Plasmid pBSm144 was produced by ligating a 4.6-kb *HindIII/NotI* fragment from the *HindIII* 'I' genomic clone of MCMV (K181

strain) between the *HindIII* and *NotI* sites of pBlueScript. The deletion plasmid pm144ZG was produced by insertion of a blunt-ended 4.6-kb *BamHI* fragment from pON855 (ref. 25) comprising a dual *lacZ/gpt* expression cassette, into the *BspEI* site of pBSm144. The cassette is thereby inserted 50 nucleotides downstream of the initiating ATG of the m144 ORF. Nucleotide numbers (in brackets) correspond to the complete genomic sequence of MCMV (Smith strain) (ref. 6) and give the position in kilobases.

Production of recombinant MCMV deleted of m144. The recombinant KΔ144 was produced by homologous recombination between K181 viral DNA and the plasmid pm144ZG. After cotransfection of viral and plasmid DNA into MEF, recombinant virus was selected by the detection of blue plaques under an X-gal overlay and cloned by three rounds of plaque purification. The revertant KΔ144r was produced by homologous recombination between KΔ144 viral DNA and plasmid pBSm144. Revertants were identified as clear plaques under an X-gal overlay and cloned by three rounds of plaque purification. The correct insertion of the *lacZ/gpt* cassette within the m144 region of KΔ144 was confirmed by Southern blotting.

Quantification of virus replication *in vitro*. MEF cultures were infected at a multiplicity of infection of 0.01 with either wild type, KΔ144 or KΔ144r overlaid with 1 ml medium per culture, and the virus was quantified at the time points indicated by plaque assay on MEF.

Quantification of virus replication *in vivo*. Female mice six to seven weeks old (obtained from the Animal Resources Centre, Murdoch, Western Australia) were used. The doses of each of the viruses used was 10⁴ plaque-forming units (PFU) for BALB/c mice, and 10^{4.3} PFU for BALB.B6-Cmv1^{-/-} mice. All virus challenges were administered intraperitoneally, and organs were collected on the days indicated, before being homogenized at 4°C in medium containing 2% (v/v) FCS, clarified by centrifugation at 2,000g and the supernatant stored at -70°C. Virus was quantified by plaque assay on MEF. The results are expressed as the log₁₀ mean PFU per organ $\pm$ s.d.; 4–6 mice were used per group.

***In vivo* depletion of cellular subsets.** First, BALB/c mice. For depletion of asialo-GM₁⁺ NK cells, BALB/c female mice were inoculated intraperitoneally at day -2 with 1 mg anti-asialo GM₁ (Wako Pure Chemicals, Japan). Both depleted and undepleted mice were challenged with 10⁴ PFU of either wild-type K181 or KΔ144 on day 0. NK cell cytotoxicity was assessed from splenic cultures at day 2 post-infection. Uninfected, undepleted mice were also included as controls. Splenocytes were incubated with ⁵¹Cr-labelled YAC-1 target cells at the indicated effector:target cell ratios (in triplicate) in a standard 4-h cytotoxicity assay. Splenic cultures from each mouse were assayed individually for NK cytotoxicity. The *in vivo* depletions of CD4⁺ and CD8⁺ T cells were performed by intraperitoneal inoculation with both YTS 169.4 (ref. 26) and YTS 191.1 (ref. 26) monoclonal antibodies on days -2 and 0. Confirmation of T-cell depletion was determined on day 2 by FACS analysis, using antibodies 3.168.8 (anti-CD8, ref. 27) and RL-172 (anti-CD4, ref. 28). Binding of the anti-T cell subset 3.168.8 and RL-172 was not inhibited by the presence of YTS-169.4 or YTS-191.1 antibodies, respectively. Second, BALB.B6-Cmv1^{-/-} mice. *In vivo* depletion of NK1.1-positive cells was performed by intraperitoneal inoculations of 6-week-old mice with monoclonal antibody PK136 on days -2 and 0, as described^{13,14}.

Received 28 September 1996; accepted 3 February 1997.

1. Banks, T. A. & Rouse, B. T. Herpesviruses-immune escape artists? *Clin. Infect. Dis.* **14**, 933–941 (1992).
2. Spriggs, M. K. One step ahead of the game: viral immunomodulatory molecules. *Annu. Rev. Immunol.* **14**, 101–130 (1996).
3. Davis-Poynter, N. J. & Farrell, H. E. Masters of deception: a review of herpesvirus immune evasion strategies. *Immunol. Cell Biol.* **74**, 512–522 (1996).
4. Beck, S. & Barrell, B. G. Human cytomegalovirus encodes a glycoprotein homologous to MHC class-I antigens. *Nature* **331**, 269–272 (1988).
5. Wiley, D. Molecular immunology: MHC gene in cytomegalovirus. *Nature* **331**, 209–210 (1988).
6. Rawlinson, W. D., Farrell, H. E. & Barrell, B. G. Analysis of the complete DNA sequence of murine cytomegalovirus. *J. Virol.* **70**, 8833–8849 (1996).
7. Bjorkman, P. J., Saper, M. A. & Wiley, D. C. Structure of human class I histocompatibility antigen, HLA-A2. *Nature* **329**, 506–512 (1987).
8. Bjorkman, P. J. *et al.* The foreign antigen binding site and T cell recognition regions of class I histocompatibility antigens. *Nature* **329**, 512–518 (1987).
9. Browne, H., Churcher, M. & Minson, T. Construction and characterization of a human cytomegalovirus mutant with the UL18 (class I homologue) gene deleted. *J. Virol.* **66**, 6784–6787 (1992).
10. Bancroft, G. J., Shellam, G. R. & Chalmers, J. E. Genetic influences on the augmentation of natural killer (NK) cells during murine cytomegalovirus infection: correlation with patterns of resistance. *J. Immunol.* **126**, 988–994 (1981).
11. Bukowski, J. F., Woda, B. A. & Welsh, R. M. Pathogenesis of murine cytomegalovirus infection in natural killer cell-depleted mice. *J. Virol.* **52**, 119–128 (1984).
12. Shellam, G. R., Flexman, J. P., Farrell, H. E. & Papadimitriou, J. M. The genetic background

- modulates the effect of the beige gene on susceptibility to cytomegalovirus infection in mice. *Scand. J. Immunol.* **22**, 147–155 (1985).
13. Scalzo, A. A. *et al.* The effect of the *Cmv1* resistance gene, which is linked to the natural killer cell gene complex, is mediated by natural killer cells. *J. Immunol.* **149**, 581–589 (1992).
 14. Scalzo, A. A. *et al.* The BALB.B6-*Cmv1*⁺ mouse: a strain congenic for *Cmv1* and the NK gene complex. *Immunogenetics* **41**, 148–151 (1995).
 15. Karlhofer, F. M. & Yokoyama, W. M. Stimulation of murine natural killer (NK) cells by a monoclonal antibody specific for the NK1.1 antigen. IL-2-activated NK cells possess additional specific stimulation pathways. *J. Immunol.* **146**, 3662–3673 (1991).
 16. Yokoyama, W. M. Natural killer cell receptors. *Curr. Opin. Immunol.* **7**, 110–120 (1995).
 17. Gumperz, J. E. & Parham, P. The enigma of the natural killer cell. *Nature* **378**, 245–248 (1995).
 18. Raulet, D. H. Recognition events that inhibit and activate natural killer cells. *Curr. Opin. Immunol.* **8**, 372–377 (1996).
 19. Kärre, K. Express yourself or die: peptides, MHC molecules, and NK cells. *Science* **267**, 978–979 (1995).
 20. Wiertz, E. J. H. J. *et al.* The human cytomegalovirus US11 gene product dislocates MHC class I heavy chains from the endoplasmic reticulum to the cytosol. *Cell* **84**, 769–779 (1996).
 21. del Val, M. *et al.* Cytomegalovirus prevents antigen presentation by blocking the transport of peptide-loaded major histocompatibility complex class I molecules into the medial-Golgi compartment. *J. Exp. Med.* **176**, 729–738 (1992).
 22. Thäle, R. *et al.* Identification of the mouse cytomegalovirus genomic region affecting MHC class I molecule transport. *J. Virol.* **69**, 6098–6105 (1995).
 23. Browne, H., Smith, G., Beck, S. & Minson, T. A complex between the MHC class I homologue encoded by human cytomegalovirus and β_2 microglobulin. *Nature* **347**, 770–771 (1990).
 24. Fahnestock, M. L. *et al.* The MHC class I homologue encoded by human cytomegalovirus binds endogenous peptides. *Immunology* **3**, 538–590 (1995).
 25. Vieira, J., Farrell, H. E., Rawlinson, W. D. & Mocarski, E. S. Genes in the Hind III J fragment of the murine cytomegalovirus genome are dispensable for growth in cultured cells: insertion mutagenesis with a *lacZ/gpt* cassette. *J. Virol.* **68**, 4837–4846 (1994).
 26. Cobbold, S. P. *et al.* Therapy with monoclonal antibodies by elimination of T cell subsets *in vivo*. *Nature* **312**, 548–550 (1984).
 27. Sarmiento, M., Glasebrook, A. L. & Fitch, F. W. IgG or IgM monoclonal antibodies reactive with different determinants on the molecular complex bearing Lyt2 antigen block T cell-mediated cytotoxicity in the absence of complement. *J. Immunol.* **125**, 2665–2672 (1985).
 28. Ceredig, R., Lowenthal, J. W., Nabholz, M. & MacDonald, H. R. Expression of interleukin-2 receptors as a differentiation marker on intrathymic stem cells. *Nature* **314**, 98–100 (1985).
 29. Pearson, W. R. & Lipman, D. J. Improved tools for biological sequence comparisons. *Proc. Natl Acad. Sci. USA* **85**, 2444–2448 (1988).
 30. Higgins, D. G., Bleasby, A. J. & Fuchs, R. Clustal V: improved software for multiple sequence alignment. *Comput. Adv. Biol. Sci.* **8**, 189–191 (1992).

Acknowledgements. We thank J. Vieira for providing plasmid pON855 and J. Allan for providing monoclonal antibodies to T-cell markers and for critical reading of the manuscript. This work was supported by the NH&MRC and AMRAD Corporation of Australia. H.E.F. is supported by the Raine Foundation, University of Western Australia.

Correspondence and requests for materials should be addressed to H.E.F. (e-mail: hef@uniwa.uwa.edu.au).

The class I MHC homologue of human cytomegalovirus inhibits attack by natural killer cells

Hugh T. Reyburn, Ofer Mandelboim, Mar Valés-Gómez, Daniel M. Davis, Laszlo Pazmany & Jack L. Strominger

Department of Molecular and Cellular Biology, Harvard University, 7 Divinity Avenue, Cambridge, Massachusetts 02138, USA

Recognition and destruction of virus-infected cells by class I major histocompatibility complex (MHC) restricted cytotoxic T lymphocytes (CTL) is a central part of the immune system's attempts to control and eliminate virus infection. It is therefore not surprising that many viruses have evolved strategies to interfere with the processing and presentation of peptide antigen on class I MHC molecules (reviewed in ref. 1). These mechanisms act to prevent or reduce expression of MHC molecules at the cell surface. However, many natural killer (NK) cells are able to recognize and destroy host cells that no longer express class I MHC molecules (the 'missing self' hypothesis²). Thus, any virus-infected cell that has lost cell-surface expression of MHC class I to avoid CTL attack should become susceptible to NK-cell-mediated destruction. We describe here the first example, to our knowledge, of a viral strategy to evade immune surveillance by NK cells.

Human cytomegalovirus (HCMV) acts to reduce cell-surface expression of MHC class I molecules by a variety of mechanisms that utilize a number of HCMV gene products, including US2, US3,

US6 and US11 (refs 3–9), and so would be expected to become susceptible to NK cell attack. Strikingly, however, HCMV encodes a β_2 -microglobulin (β_2m)-binding protein with homology to MHC class I heavy chains^{10,11} and which is able to bind peptides¹². This MHC homologue (the UL18 gene product) might interfere with some aspect of the cell-mediated immune response of the host¹³, possibly acting as a surrogate MHC class I molecule able to engage NK-cell inhibitory receptors (NKIR) and so avoid attack by NK cells¹².

To test whether expression of the UL18 protein could protect from NK cell attack, we transfected the UL18 gene into the HLA-A, -B, -C-negative human B-cell line 721.221¹⁴ and assayed the ability of NK cells to kill this transfectant. Surface expression of the UL18 gene product was assessed by flow cytometric analysis (Fig. 1) using the β_2m -specific monoclonal antibody BBM1 (ref. 15). As the staining of the 721.221/UL18 transfectant with BBM1 was low, expression of UL18 on the cell surface was confirmed by surface labelling using biotin¹⁶ and immunoprecipitation (Fig. 2). The higher relative molecular mass (M_r) of the UL18 gene product¹¹ is due to its heavy glycosylation (see below). The fact that the N-linked glycans of the UL18 protein are resistant to digestion by Endo H indicates that this molecule has passed through the Golgi, an indicator of surface expression. The low level of expression of the UL18 transfectant (Fig. 1), compared to that of the classical class I molecule HLA-Cw6, was observed in all of the transfectants analysed. Moreover, for reasons that are not clear, all transfectants made so far lost expression of UL18 after some time in culture. Modifications of the gene designed to enhance transcription and translation did not increase the low level of expression, neither did inclusion of RNA-stabilizing sequences from the β -globin gene inserted 5' and 3' of the UL18 gene product. This loss of expression occurred with transfectants prepared with either integrating or episomal vectors. Accordingly, cell-surface expression of UL18 was checked on the day of each experiment.

Despite its poor surface expression, the UL18 transfectant could be as potent at inhibiting lysis by NK cells as classical class I MHC molecules (Fig. 3). Further, and in contrast to the HLA-Cw6 and Cw7 transfectants which inhibited NK cells in a group-specific manner, the 721.221/UL18 transfectant was protected from lysis by both NK1- and NK2-specific NK cells (Fig. 3). This protection was reproducible with NK lines and clones generated from several donors and with many independently derived UL18-expressing transfectants. Transfectants that had lost surface expression of UL18 were not protected from lysis by NK cells (data not shown).

The ability of UL18 to inhibit lysis of 721.221 cells by both NK1- and NK2-specific NK cells was surprising. Two possible mechanisms were considered. First, the UL18 molecule might interact with NK inhibitory receptors representing all the defined specificities; second, the inhibitory signal may be delivered to the NK cell by some other widely expressed NK cell-surface molecule.

The UL18 gene product is heavily glycosylated, containing 13 potential N-linked glycosylation sites (about eight of which seem to be used) which account for the different M_r values of UL18 and the HLA heavy chain (Fig. 2)^{10,11}. As many of the cell-surface molecules preferentially expressed on NK cells are C-type lectins, we used various glycosidase inhibitors to modify the composition of N-linked glycans on the UL18 gene product before testing the ability of the UL18 transfectant to inhibit NK cells. Treatment with either deoxynojirimycin (dNM, an inhibitor of endoplasmic reticulum glucosidases I and II) or deoxymannojirimycin (dMM, an inhibitor of Golgi mannosidase I) results in the expression of proteins bearing high-mannose glycans rather than complex sialated sugars¹⁷. Neither dNM nor dMM treatment reduced the surface expression of the UL18 gene product; indeed, dMM treatment seemed to increase the expression of UL18 (Fig. 4 legend). However, treatment with dNM or dMM markedly reduced or prevented protection by UL18 against NK lysis (Fig. 4). Treatment of target cells with these

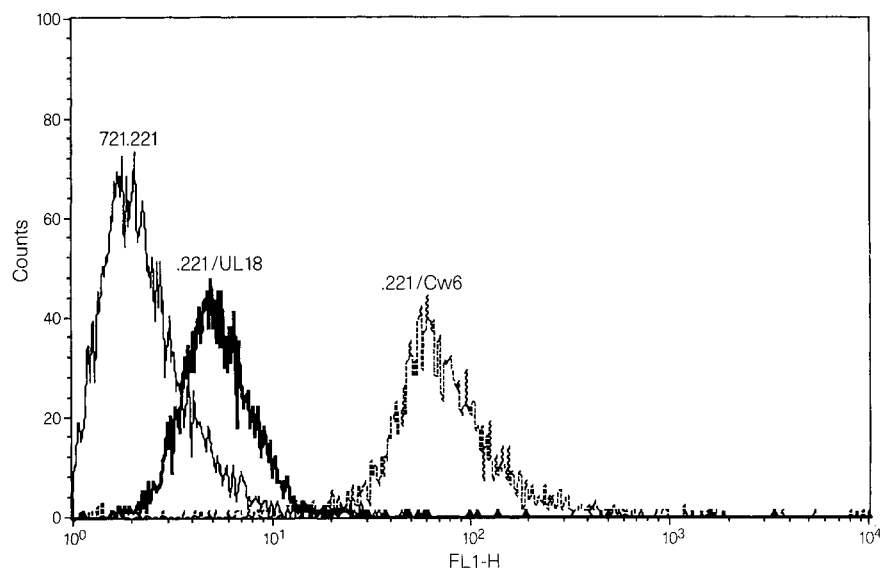


Figure 1 Expression of UL18 and of HLA-Cw6 on transfected 721.221 cells. Analysis of expression of β 2m, and so indirectly both the UL18 gene product and HLA-Cw6, by FACS analysis using the BBM1 monoclonal antibody. Mean fluorescence intensities were: control 721.221 cells, 2.21; 721.221/UL18, 6.10 and 721.221/Cw6, 79.8.

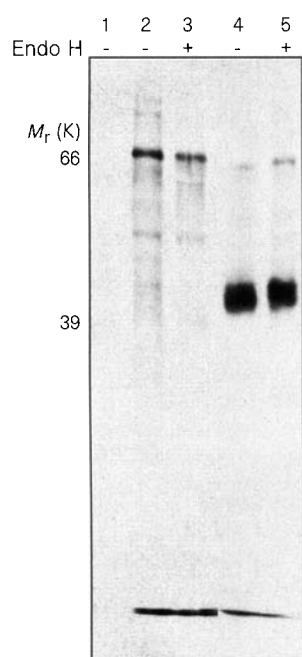


Figure 2 The UL18 gene product is expressed at the cell surface. 721.221, .221/UL18 and .221/Cw6 were surface-labelled with biotin, immunoprecipitated with mAb BBM1 and biotinylated proteins visualized with avidin-ECL after SDS-PAGE and immunoblotting. As previously reported²⁶, untransfected 721.221 cells have little or no surface expression of β 2m (lane 1). A ~67K molecule, the reported M_r for UL18¹⁸, was precipitated from the UL18 transfectant (lane 2); a 44K molecule, the heavy chain of HLA-Cw6, was precipitated from .221/Cw6 (lane 4, both in association with the 12K β 2m (lower band). Both the UL18 and Cw6 heavy chains were resistant to Endo H digestion (lanes 3 and 5). The migration positions of markers are shown.

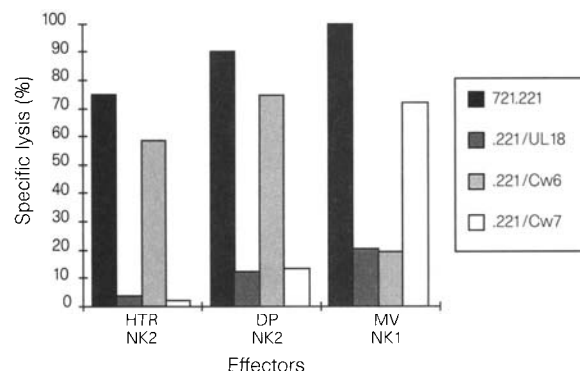


Figure 3 The UL18 gene product confers protection from lysis by both NK1 and NK2 lines (effector-to-target ratio (E:T) was 3:1; HLA types of donors: HTR (A1, A2; B7, B8: Cw7), DP (A1, A3; B7, B8: Cw7), MV (A2, A11; B18, B44; Cw5).

glycosidase inhibitors inhibited lysis by NK cells, presumably as a result of an effect on other molecules involved in lysis. In contrast, the specific inhibition of NK lysis associated with expression of either HLA-Cw3 or Cw4 was unaffected by carbohydrate modification. Essentially identical data were obtained from analysis of NK lines (Fig. 4) and of NK clones (not shown). Thus correct glycosylation of the UL18 molecule, although not required for surface expression or association with β 2m, is critically important for mediating inhibition of NK cell lysis. There are only a few examples in which inhibiting conversion of high-mannose to complex-type glycans has any biological consequence. In view of the importance of lectin-mediated recognition in NK killing, and the extensive glycosylation of UL18, these results underscore the importance of N-linked glycans in cellular recognition.

NK cells express a number of C-type lectins at the cell surface, one of which may be a receptor for UL18. A plausible candidate for the lectin receptor was CD94, a C-type lectin expressed on most NK cells and on a subset of T lymphocytes^{18,19}. Antibody-mediated ligation of this molecule may inhibit, augment, or have no effect on NK cell killing²⁰. CD94 may be expressed at the cell surface alone or

dimerized with other proteins^{21,22}. Moreover, a complex of CD94 and NKG2 can recognize MHC molecules and inhibit NK killing²³. Inclusion in the killing assay of monoclonal antibody HP-3D9, which is specific for CD94, abolished the target-cell protection mediated by UL18 (Fig. 5). Inhibition mediated by HLA-Cw6 and -Cw7 was not reversed by an antibody specific for CD94, although it was appropriately reversed by monoclonal antibodies HP3E4 and GL183, which recognize NKIR1 and NKIR2 respectively. On the other hand, inclusion of saturating amounts of NKIR-specific antibodies in the killing assays, enough to reverse protection conferred by HLA-C molecules, had no effect on the inhibition of killing by UL18 (Fig. 5). These results indicate that CD94 is a component of the UL18 receptor and that the NKIRs detected by GL183 and HP3E4, at least, are not involved.

We have described a new viral strategy for evading immune surveillance. HCMV seems to have evolved two complementary ways to avoid immune recognition of the cell it infects. An HCMV gene product that blocks TAP function⁹ has been identified as US6 (H. Ploegh, personal communication). Expression of the US3, US2 and US11 gene products leads to retention of host class I MHC

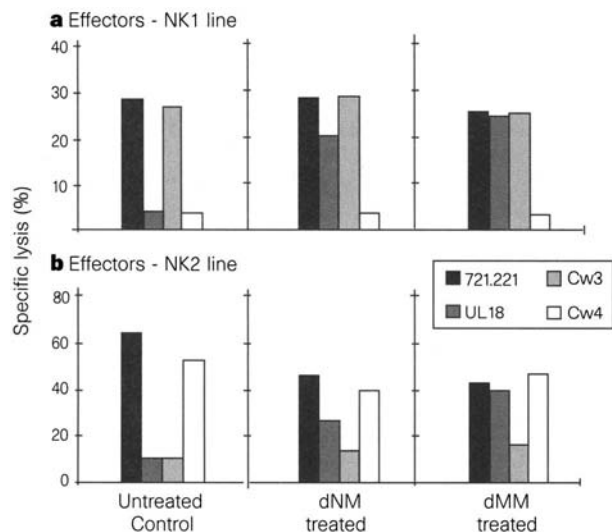


Figure 4 Effect of glycosidase inhibitors on lysis of 721.221 cells transfected with UL18, HLA-Cw3 or -Cw4. After treatment (see Methods), mean fluorescence intensities were: untreated 721.221/UL18 transfectant stained with isotype-matched control antibody, 2.33, and with BBM1, 4.99; dNM treated cells: isotype control, 2.29, and BBM1, 4.02; dMM treated cells: isotype control, 2.62, and BBM1, 10.37.

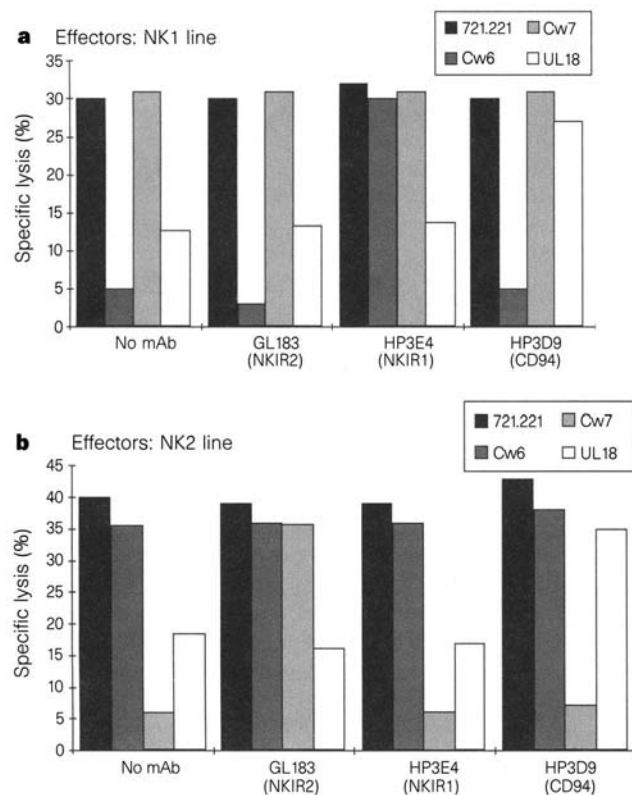


Figure 5 UL18-mediated inhibition of lysis by **a**, NK1, and **b**, NK2 specific NK cells (NK lines; E:T = 3:1) is reversed by CD94-specific mAb, but not by NKIR-specific mAb. The antibodies were used at $10 \mu\text{g ml}^{-1}$.

molecules in the endoplasmic reticulum and, later in infection, specific proteolysis of these molecules³⁻⁸, thereby blocking CTL recognition and killing of the infected cell. Expression of UL18 leads to inhibition of NK cell attack. HCMV may have had to acquire such defences because of its prolonged replication period, with an extended exposure to immune attack. If a virus that has co-evolved with humans for millions of years²⁴ has evolved a strategy to evade NK surveillance, the need to avoid NK cell activity must be a strong selective pressure on pathogen evolution.

The role of a C-type lectin as an inhibitory receptor on human NK cells is also notable. In the mouse, the principal inhibitory receptors are the C-type lectin Ly-49 molecules, whereas in humans the previously identified NKIRs are members of the immunoglobulin superfamily. These separate types of molecule serve the same inhibitory function. The existence of a second type of inhibitory receptor in human NK cells, CD94, a C-type lectin^{20,23}, indicates that multiple mechanisms have evolved to protect host cells against NK cells, at least some of which have been exploited by pathogens to evade immune surveillance.

Methods

721.221 cells were obtained from the ATCC. The various HLA-C transfectants have been described²⁵. The UL18 gene was subcloned from p103 (ref. 11) into pCDNA3 (Invitrogen) and transfected into 721.221 cells by electroporation. Transfectants were selected in 1.6 mg ml^{-1} Geneticin (Gibco BRL) and screened for surface expression of $\beta 2\text{m}$ by FACS analysis and immunoprecipitation using the BBM1 monoclonal antibody (hybridoma obtained from the ATCC). For cell surface biotinylation, cells were washed four times with PBS, 1 mM MgCl_2 , 0.1 mM CaCl_2 and then incubated at 4°C with Sulfo-NHS Biotin (Pierce) (0.5 mg ml^{-1} in PBS, 1 mM MgCl_2 , 0.1 mM CaCl_2) for 30 min. Cells were then washed extensively to remove unbound biotin and lysed in 1% NP-40. After immunoprecipitation and SDS-PAGE, biotinylated proteins were visualized by immunoblotting and detection with avidin-ECL reagents (Amersham). Endo H was purchased from New England Biolabs and digestion was in the buffer supplied by the manufacturer for 1 h at 37°C .

Generation of NK cells and cytotoxicity assays have been described²⁵. mAb HP3E4 was a gift from M. López-Botet. mAb GL183 was from Immunotech and HP-3D9 from Pharmingen. mAbs were used in killing assays at $10 \mu\text{g ml}^{-1}$.

For treatment with dNM or dMM (Boehringer Mannheim), the various transfected and untransfected cell lines were cultured for 48 h at $5 \times 10^5 \text{ cells ml}^{-1}$ in RPMI 10% FCS alone, or supplemented with dNM or dMM (2 mM final concentration). These conditions result in $\sim 90\%$ of modified HLA proteins and 10% unmodified (as judged by immunoprecipitation, after cell-surface labelling and Endo-H digestion; data not shown). Unmodified HLA proteins result from residual HLA that were present at the start of culture and are not affected by the treatment.

Received 12 December 1996; accepted 7 March 1997.

1. Spriggs, M. K. One step ahead of the game: Viral immunomodulatory molecules. *Annu. Rev. Immunol.* **14**, 101–130 (1996).
2. Ljunggren, H.-G. & Karre, K. In search of the 'missing self': MHC molecules and NK recognition. *Immunol. Today* **11**, 237–242 (1990).
3. Beersma, M. F. C., Bijlmakers, M. J. E. & Ploegh, H. L. Human cytomegalovirus down-regulates HLA class I expression by reducing the stability of class I H chains. *J. Immunol.* **151**, 4455–4464 (1993).
4. Jones, T. R. *et al.* Multiple independent loci within the human cytomegalovirus unique short region down-regulate expression of major histocompatibility complex class I heavy chains. *J. Virol.* **69**, 4830–4841 (1995).
5. Wiertz, E. J. H. J. *et al.* The human cytomegalovirus US11 gene product dislocates MHC class I heavy chains from the endoplasmic reticulum to the cytosol. *Cell* **84**, 769–779 (1996).
6. Wiertz, E. J. H. J. *et al.* Cytosolic destruction of a type I membrane protein by transfer via the Sec61p complex from the ER to the proteasome. *Nature* **384**, 432–438 (1996).
7. Ahn, K. *et al.* Human cytomegalovirus inhibits antigen presentation by a sequential multistep process. *Proc. Natl Acad. Sci. USA* **93**, 10990–10995 (1996).
8. Jones, T. R. *et al.* Human cytomegalovirus US3 impairs transport and maturation of major histocompatibility complex class I heavy chains. *Proc. Natl Acad. Sci. USA* **93**, 11327–11333 (1996).
9. Hengel, H., Flohr, T., Hammerling, G. J. & Koszinowski, U. H. Human cytomegalovirus inhibits peptide translocation into the endoplasmic reticulum for MHC class I assembly. *J. Gen. Virol.* **77**, 2287–2296 (1996).
10. Beck, S. & Barrell, B. G. Human cytomegalovirus encodes a glycoprotein homologous to MHC class-I antigens. *Nature* **331**, 269–272 (1988).
11. Browne, H., Smith, G., Beck, S. & Minson, A. A complex between the MHC class I homologue encoded by human cytomegalovirus and β_2 -microglobulin. *Nature* **347**, 770–772 (1990).
12. Fahnestock, M. L. *et al.* The MHC class I homologue encoded by human cytomegalovirus binds endogenous peptides. *Immunity* **3**, 583–590 (1995).
13. Wiley, D. MHC gene in cytomegalovirus. *Nature* **331**, 209–210 (1988).

14. Shimizu, Y. & DeMars, R. Production of human cells expressing individual transferred HLA-A, -B, -C genes using an HLA-A, -B, -C null human cell line. *J. Immunol.* **142**, 3320 (1989).
15. Parham, P., Androlewicz, M. J., Holmes, N. J. & Rothenberg, B. E. Arginine-45 is a major part of the determinant of human $\beta 2$ -microglobulin recognised by mouse monoclonal antibody BBM.1. *J. Biol. Chem.* **258**, 6179–6186 (1983).
16. Brocker, T., Peter, A., Trautwein, A. & Karjalainen, K. New simplified molecular design for functional T cell receptor. *Eur. J. Immunol.* **23**, 1435–1439 (1993).
17. Fuhrmann, U., Bause, E. & Ploegh, H. Inhibitors of oligosaccharide processing. *Biochim. Biophys. Acta* **825**, 95–110 (1985).
18. Aramburu, J. *et al.* A novel functional cell surface dimer (Kp43) expressed by natural killer cells and T cell receptor-g/d+ T lymphocytes. I. Inhibition of the IL-2 dependent proliferation by anti-Kp43 monoclonal antibody. *J. Immunol.* **144**, 3238–3247 (1990).
19. Aramburu, J., Balboa, M. A., Izquierdo, M. & Lopez-botet, M. A novel functional cell surface dimer (Kp43) expressed by natural killer cells and T cell receptor-g/d+ T lymphocytes. II. Modulation of natural killer cell cytotoxicity by anti-Kp43 monoclonal antibody. *J. Immunol.* **147**, 714–721 (1991).
20. Perez-Villar, J. J. *et al.* Functional ambivalence of the Kp43 (CD94) NK cell associated surface antigen. *J. Immunol.* **154**, 5779–5788 (1995).
21. Phillips, J. H. *et al.* CD94 and a novel associated protein (94AP) form a NK cell receptor involved in the recognition of HLA-A, HLA-B and HLA-C allotypes. *Immunity* **5**, 163–172 (1996).
22. Perez-Villar, J. J. *et al.* Biochemical and serologic evidence for the existence of functionally distinct forms of the CD94 NK cell receptor. *J. Immunol.* **157**, 5367–5374 (1996).
23. Lazetic, S., Chang, C., Houchins, J. P., Lanier, L. L. & Phillips, J. H. Human natural killer cell receptors involved in MHC class I recognition are disulfide linked heterodimers of CD94 and NKG2 subunits. *J. Immunol.* **157**, 4741–4745 (1996).
24. McGeoch, D. J., Cook, S., Dolan, A., Jamieson, F. E. & Telford, E. A. Molecular phylogeny and evolutionary timescale for the family of mammalian herpesviruses. *J. Mol. Biol.* **247**, 443–458 (1995).
25. Mandelboim, O. *et al.* Protection from lysis by natural killer cells of group 1 and 2 specificity is mediated by residue 80 in human histocompatibility leukocyte antigen C alleles and also occurs with empty major histocompatibility complex molecules. *J. Exp. Med.* **184**, 913–922 (1996).
26. Shimizu, Y., Geraghty, D. E., Koller, B. H., Orr, H. T. & DeMars, R. Transfer and expression of three cloned human non-HLA-A, B, C class I major histocompatibility genes in mutant lymphoblastoid cells. *Proc. Natl Acad. Sci. USA* **85**, 227–231 (1988).

Acknowledgements. We thank H. Browne, for the UL18 gene, M. López-Botet for mAb HP3E4, and M. L. Hedley for discussion. This work was supported by fellowships to H.T.R. (Wellcome Trust), O.M. (EMBO and Fulbright) and L.P. (Arthritis Foundation) and by an NIH grant.

Correspondence and requests for materials should be addressed to J.L.S. (e-mail: jlstrom@fas.harvard.edu).

Viral FLICE-inhibitory proteins (FLIPs) prevent apoptosis induced by death receptors

Margot Thome*, Pascal Schneider*, Kay Hofmann†, Helmut Fickenscher†, Edgar Meinl‡, Frank Neipel‡, Chantal Mattmann*, Kim Burns*, Jean-Luc Bodmer*, Michael Schröter*, Carsten Scaffidi§, Peter H. Krammer§, Marcus E. Peter§ & Jürg Tschopp*

* Institute of Biochemistry, University of Lausanne, and † Swiss Institute for Experimental Cancer Research, BIL Biomedical Research Center, Chemin des Boveresses 155, CH-1066 Epalinges, Switzerland
‡ Institute for Clinical and Molecular Virology, Friedrich-Alexander University Erlangen-Nürnberg, Schlossgarten 4, D-91054 Erlangen, Germany
§ Division of Immunogenetics, Tumor Immunology Program, German Cancer Research Center, Im Neuenheimer Feld 280, D-69120 Heidelberg, Germany

Viruses have evolved many distinct strategies to avoid the host's apoptotic response^{1,2}. Here we describe a new family of viral inhibitors (v-FLIPs) which interfere with apoptosis signalled through death receptors³ and which are present in several γ -herpesviruses (including Kaposi's-sarcoma-associated human herpesvirus-8), as well as in the tumorigenic human molluscipoxvirus⁴. v-FLIPs contain two death-effector domains which interact with the adaptor protein FADD^{5,6}, and this inhibits the recruitment and activation of the protease FLICE^{7,8} by the CD95 death receptor³. Cells expressing v-FLIPs are protected against apoptosis induced by CD95 or by the related death receptors TRAMP^{9–12} and TRAIL-R. The herpesvirus saimiri FLIP is detected late during the lytic viral replication cycle, at a time when host cells are partially protected from CD95-ligand-mediated apoptosis. Protection of virus-infected cells against death-receptor-induced apoptosis may lead to higher virus pro-

duction and contribute to the persistence and oncogenicity¹³ of several FLIP-encoding viruses.

In the immune system, apoptosis is used as a defence strategy to eliminate potentially harmful agents. For example, virus-specific T lymphocytes kill infected cells by delivering pro-apoptotic proteins (perforin and granzymes) or by expressing CD95 (Apo-1/Fas) ligand and triggering cell death of target cells via CD95 (ref. 14). Alternatively, apoptosis of infected cells may be cell-autonomous; cells detect the presence of certain viruses and respond by committing suicide. To allow increased virus replication, some viruses carry genes whose products are able to interfere with the host's apoptotic machinery¹. There is increasing evidence that apoptotic cell death requires the activation of members of the ICE-like family of cysteine proteases (caspases)¹⁵ and, not surprisingly, apoptosis inhibitory proteins that block these effector proteases are found in some viruses—for example, CrmA in the cowpox virus or p35 in baculovirus¹. Other viral anti-apoptotic genes resemble the mammalian *bcl-2* gene¹.

The most direct pathway leading to the activation of ICE-like proteases and thus to cell death appears to be triggered by some members of the tumour-necrosis-factor receptor (TNF-R) superfamily, CD95, TNFR-1 and TRAMP (ws1/DR-3/Apo-3)^{3,9–12} (collectively called death receptors). These receptors relay death signals through a cytoplasmic sequence motif called the death domain (DD), which interacts with the DD of the adaptor molecules FADD and/or TRADD³, recruiting them to the membrane. FADD then associates with the ICE-like protease FLICE (caspase-8, Mch5, MACH)^{7,8} through death-effector domains (DEDs) present at the carboxy terminus of FADD and the amino terminus of FLICE, leading to the assembly of a receptor-associated death-inducing signalling complex (DISC)¹⁶. DISC-associated FLICE subsequently initiates proteolytic activation of other ICE family members, which in turn leads to apoptosis^{7,8,17}.

Deletion mutants of FADD (containing only the DD) or FLICE (containing two DEDs) can act as dominant-negative inhibitors^{8,18} early in the signalling pathway of death receptors, raising the possibility that structurally related natural inhibitors of apoptosis may exist. We therefore screened public databases with a generalized profile¹⁹ constructed from DEDs of human and murine FADD, FLICE and Mch4 (ref. 20; caspase-10). Using this profile, the ORF E8 within the genome of equine herpesvirus-2 (EHV-2) was found to contain two sequence motifs with highly significant homology to DEDs ($P < 10^{-2}$). E8 has ~23% sequence identity to the N-terminal part of FLICE and Mch4, and the presence of two copies of bona fide DEDs is evident from the selective conservation of a specific set of amino acids. EHV-2 is a γ -herpesvirus, and subsequent database rescans revealed that other γ -herpesviruses, that is, bovine herpesvirus-4, herpesvirus saimiri (HVS), human herpesvirus-8 (HHV-8) and the human molluscipoxvirus (molluscum contagiosum virus, MCV), also contain ORFs predicted to code for proteins consisting of two DED motifs (Fig. 1). We call these proteins v-FLIPs (for viral FLICE-inhibitory proteins; see below).

Considering the homology between the two DEDs of the v-FLIPs and the two DEDs found at the N terminus of FLICE, we reasoned that FLIP members may bind to FADD or a FADD-like molecule, thereby interfering with death receptor signalling. To test this hypothesis, 293T cells were transfected with expression vectors encoding FADD and N-terminally Flag-tagged FLIP from EHV-2 (Flag-E8). It was found that the 23K E8-FLIP did indeed interact with FADD in coimmunoprecipitation experiments (Fig. 2a). Moreover, when E8 and FADD were cotransfected with a Myc-tagged CD95 construct (encompassing the cytoplasmic domain including the DD), FADD bound to E8-FLIP also interacted with CD95 (Fig. 2a), indicating that FADD/E8-FLIP complex formation (mediated through DEDs) still permitted the CD95/FADD interaction (through DDs) to occur. Similar results were obtained with MCV(ORF159L)-FLIP (Fig. 2b) and HVS (ORF71)-FLIP (data not

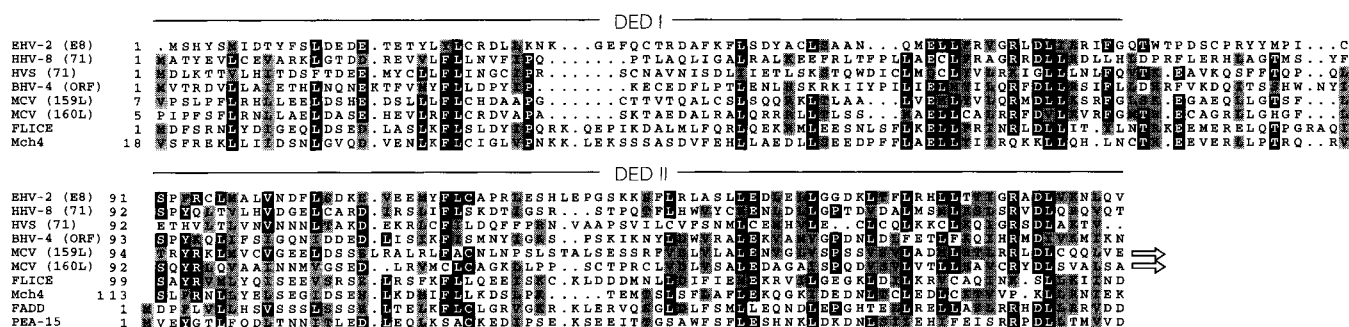


Figure 1 Amino-acid sequence alignment of the viral FLIPs and the DED containing proteins FADD, FLICE, Mch4 (GenBank (GB) accession number: U60519) and PEA-15 (GB X86809). FLIPs were found in equine herpesvirus-2 (EHV-2; GB: U20824, bovine herpesvirus-4 (BHV-4; GB: Z46385), herpesvirus saimiri (HVS; GB: H64346), human herpesvirus-8 (HHV-8)²⁹ and molluscum contagiosum virus (MCV, two distinct genes, GB: U60315). The previously published sequence of HHV-8 ORF 71 (K13)²⁹ contains a frameshift within the second DED at codon

130, thus predicting a protein of 140 amino acids. The complete sequence of 189 amino acids is shown here (GB: U90534). For each block of aligned sequences, black boxes indicate $\geq 50\%$ amino-acid sequence identity; grey shading indicates $\geq 50\%$ sequence similarity through conservative amino-acid substitutions. The arrows at the C terminus of the second DED of the MCV-FLIP sequences refer to their C-terminal extensions of 66 (159L) and 202 (160L) amino acids, respectively.

shown). Thus, v-FLIPs interact with the adaptor molecule FADD, thereby forming a complex with CD95.

As some FLIP-encoding viruses are lymphotropic, we next asked whether in the human Raji B cell line stably transfected with E8-FLIP, the viral protein would incorporate into the CD95-associated death-inducing signalling complex (DISC). To address this question, we generated Flag-E8-expressing Raji clones (RE8/11 and RE8/19) and control clones transfected with vector only (RCo/1 and RCo/3) (Fig. 2c). Two-dimensional gel electrophoresis analysis of anti-CD95 immunoprecipitates obtained from these clones revealed a 23K protein with the theoretically expected pI value of 5.0 of Flag-E8 (Fig. 2d). This protein also comigrated with Flag-E8 immunoprecipitated from E8-expressing cells, and associated with the DISC in a ligand-dependent manner (data not shown). To test whether the association of E8-FLIP with the anti-CD95 immunoprecipitates of activated cells would interfere with the recruitment of any of the DISC components CAP 1–6 (refs 16, 17), DISC formation was analysed in detail (Fig. 2e). In control RCo/3 cells, the CAP proteins CAP1 (FADD), CAP2 (hyperphosphorylated FADD), CAP3 (unknown protein containing the N terminus of FLICE), CAP4 (pro-FLICE), and CAP5 and CAP6 (the cleaved FLICE prodomain as an indicator for FLICE activation by the DISC¹⁷) were associated with CD95. In the E8-FLIP transfected RE8/19 cells however, the DISC was incomplete. Recruitment of CAP1 (FADD) was unchanged, whereas that of CAP4 (FLICE) and CAP3 was strongly reduced. Furthermore, the CAP5 and CAP6 spots were undetectable, indicating that the inhibition of apoptosis by E8-FLIP occurred at the level of FLICE recruitment and activation. FLICE is only activated when incorporated into the DISC¹⁷. Proteolytic processing by the DISC of exogenously added FLICE *in vitro* was highly decreased in the E8-FLIP transfected clone (Fig. 2f), which correlates with the reduced FLICE incorporation into the DISC (Fig. 2e).

As FLICE has been proposed to be implicated in CD95-death signalling, we anticipated that the inhibition of FLICE recruitment to the DISC should result in a decreased sensitivity to CD95 agonists. Indeed, when apoptosis was induced by anti-CD95 treatment in Raji B-cell clones, E8-FLIP-expressing clones showed a considerably decreased susceptibility to apoptosis relative to controls (Fig. 3a). E8-FLIP expression also interfered with CD95L-induced cell death in transfected Jurkat human leukaemic T cells, where the amount of E8-FLIP present correlated with the relative resistance to death (Fig. 3b). In contrast, E8-FLIP expression did not

alter the susceptibility of the clones to apoptosis induced by staurosporine or growth factor withdrawal (data not shown). In 293T human embryonic kidney cells, overexpression of CD95 led to massive cellular death, which was efficiently reduced by co-expression of E8-FLIP to an extent that is comparable to that obtained by the general ICE-like protease inhibitor z-VAD-fmk (Fig. 3c). Other v-FLIPs (MHV-159L and HVS-71) revealed similar inhibitory activities (Fig. 3d, and data not shown).

TRAMP (DR3/wsl/Apo-3) has been identified as a third member of the death receptor subfamily (containing DD) of TNF receptors^{9–12}. Like the TNFR-1, TRAMP binds FLICE through association of the adaptor proteins TRADD and FADD. Although the ligand of TRAMP remains to be identified, overexpression of the receptor itself in 293T cells induces massive cell death, which is blocked by coexpression of E8-FLIP, 71-FLIP or 159L-FLIP (Fig. 3a, f, g).

TRAIL is an orphan member of the TNF family of ligands that potently induces apoptosis in various cell lines of haematopoietic origin, including Jurkat T cells²¹. In view of the functional and structural similarity of TRAIL to CD95L, it is likely that the sequence of the uncharacterized receptor for TRAIL (TRAIL-R) and CD95 are homologous and that their death signalling pathways are similar. Indeed, the Jurkat clone with the highest level of E8-FLIP (JE8/13; Fig. 3b), incubated with increasing concentrations of recombinant soluble TRAIL (sTRAIL), was completely resistant to sTRAIL-mediated apoptosis, whereas clones JE8/1 and JE8/10 (harbouring intermediate levels of E8-FLIP) were partially protected (Fig. 3h).

Finally, we selected the HVS-71-FLIP to determine at which stages of the viral replication cycle FLIP is expressed. In a marmoset T-cell line transformed with HVS (P-1079), which produces low quantities of viral particles²², a 71-FLIP-containing mRNA of 5 kb was detected (Fig. 4a). Moreover, when fully permissive OMK cells were infected with the cytopathic HVS strains C488 or A11, 71-FLIP transcripts were detected one day before massive cellular lysis occurred (Fig. 4a). Expression of 71-FLIP coincided with increased resistance to CD95L-mediated apoptosis (Fig. 4b). Thus, 71-FLIP is expressed late in the lytic replication cycle of HVS and may protect infected cells from premature apoptosis induced by the viral overload, similar to the role proposed for other viral inhibitors such as the baculoviral p35 and IAPs, and the adenoviral Bcl-2 homologue E1B19K (ref. 1). Loss of function of these viral genes leads to a decrease in the productivity of infection of these viruses, which is accountable by an increase in apoptosis of the infected cells².

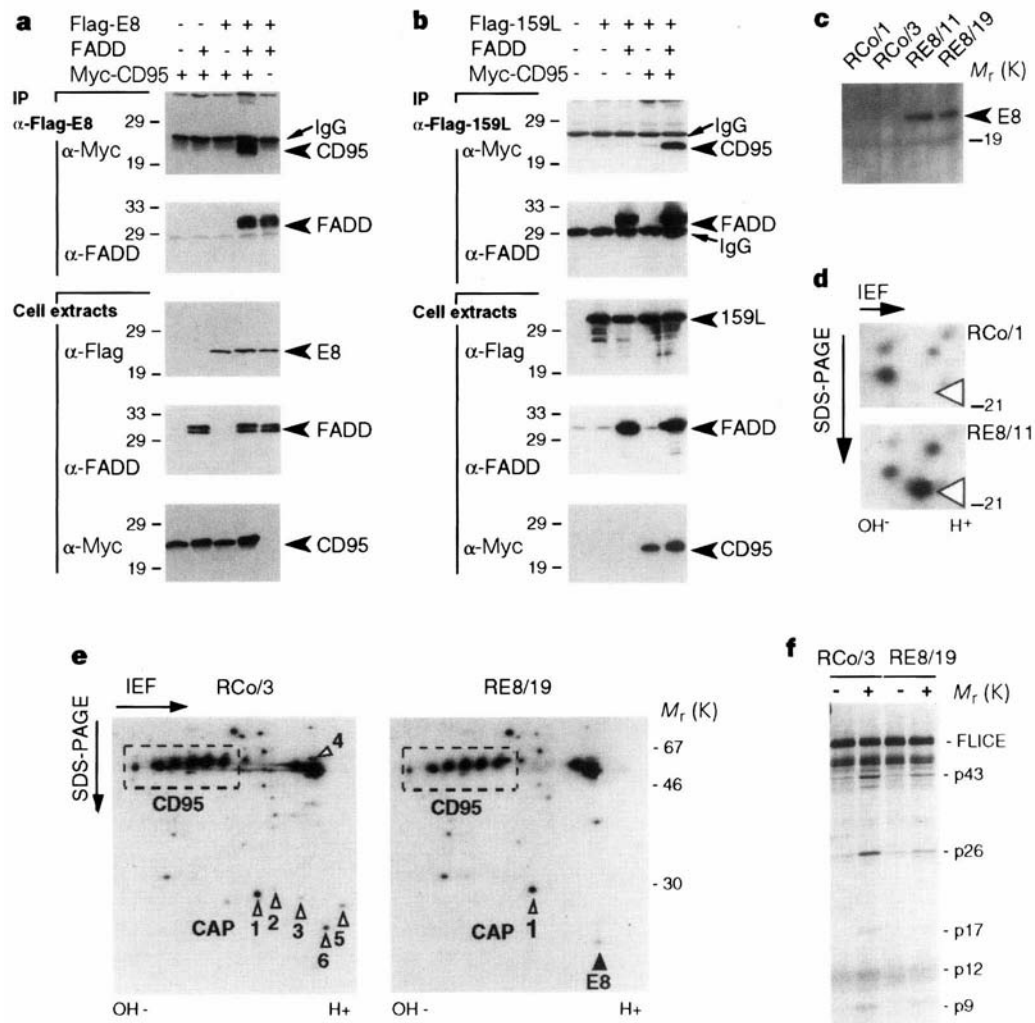


Figure 2 v-FLIPs interfere with CD95 signaling by binding to FADD. **a**, 293T cells were transiently transfected with 4 μ g of each expression vector encoding Flag-E8-FLIP, FADD and Myc-CD95 as indicated, or empty vector to keep the total amount of plasmid constant. Cells were lysed 30 h after transfection, and anti-Flag immunoprecipitates or total cell extracts were analysed for the presence of FADD, Flag-E8 or Myc-CD95 by western blotting as indicated. **b**, The association of MCV-159L-FLIP with FADD and CD95 was analysed as described in **a**. **c**, Flag-E8-FLIP expressing Raji clones RE8/11 and RE8/19, as well as the control clones RCo/1 and RCo/3 were tested for E8 expression by western blotting using an anti-Flag antibody. **d**, Analysis of anti-Apo-1 immunoprecipitates from ³⁵S-labelled control (upper panel) or Flag-E8-FLIP-expressing (lower panel) Raji clones by 2D gel electrophoresis and autoradiography. The arrow indicates the migration

position of a 23K protein of pI 5.0 found in the E8-FLIP-expressing clone RE8/11 and absent in the control clone RCo/1. **e**, Apo-1 immunoprecipitates from ³⁵S-labelled, anti-Apo-1-treated clones RCo/3 and RE8/19 were analysed by 2D gel electrophoresis and autoradiography. Numbers refer to CAP proteins: CAP1 and CAP2 (FADD), CAP3, CAP4 (FLICE), CAP5 and CAP6 (the prodomains of FLICE as a result of proteolytic cleavage). A filled arrowhead indicates the position of E8. **f**, FLICE-processing activity associated with the DISC in E8 expressing and control Raji cells. Cells were treated with anti-Apo-1 for 5 min, and Apo-1 immunoprecipitates from treated (+) or untreated (−) cells were analysed for FLICE processing activity by incubation with *in vitro* translated ³⁵S-labelled FLICE. FLICE-specific cleavage products (p43, p26, p17, p12 and p9)¹⁷ were detected by SDS-PAGE and autoradiography.

Interestingly, all FLIP encoding γ -herpesviruses also have a Bcl-2 homologue. The anti-apoptotic Bcl-2 family members block cell death induced by growth factor deprivation, γ -irradiation and cytotoxic drugs^{23,24}. However, in contrast to the v-FLIPs, these proteins have a less potent effect on CD95-mediated apoptosis of lymphoid cell lines. Some viruses may thus take advantage of two complementary anti-apoptotic functions provided on the one hand by a Bcl-2 homologue and on the other by v-FLIP.

Tissue homeostasis is maintained through the delicate balance of cell growth and apoptosis. Recent evidence indicates that tumour cells such as melanomas and hepatomas do not respond to CD95L-mediated apoptosis owing either to the downregulation of CD95 expression or to a blockade in the CD95 signalling pathway^{25,26}.

Thus, by interference with CD95 and other death-receptor-mediated signals, we propose that v-FLIPs may not only facilitate viral spread and persistence, but that they also contribute to the transforming capacity of some herpesviruses. This possibility is supported by the fact that MCV produces slow-growing epidermal neoplasms that persist for long periods with little immune response. HVS causes tumours in New World primates, and tight epidemiological links indicate that HHV-8 is an infectious cofactor for Kaposi's sarcoma and primary effusion lymphoma¹³. Future investigations should address the possible presence of v-FLIPs in tumours caused by viruses.

Note added in proof: Related findings have been reported by J. Bertin *et al.*³⁰.

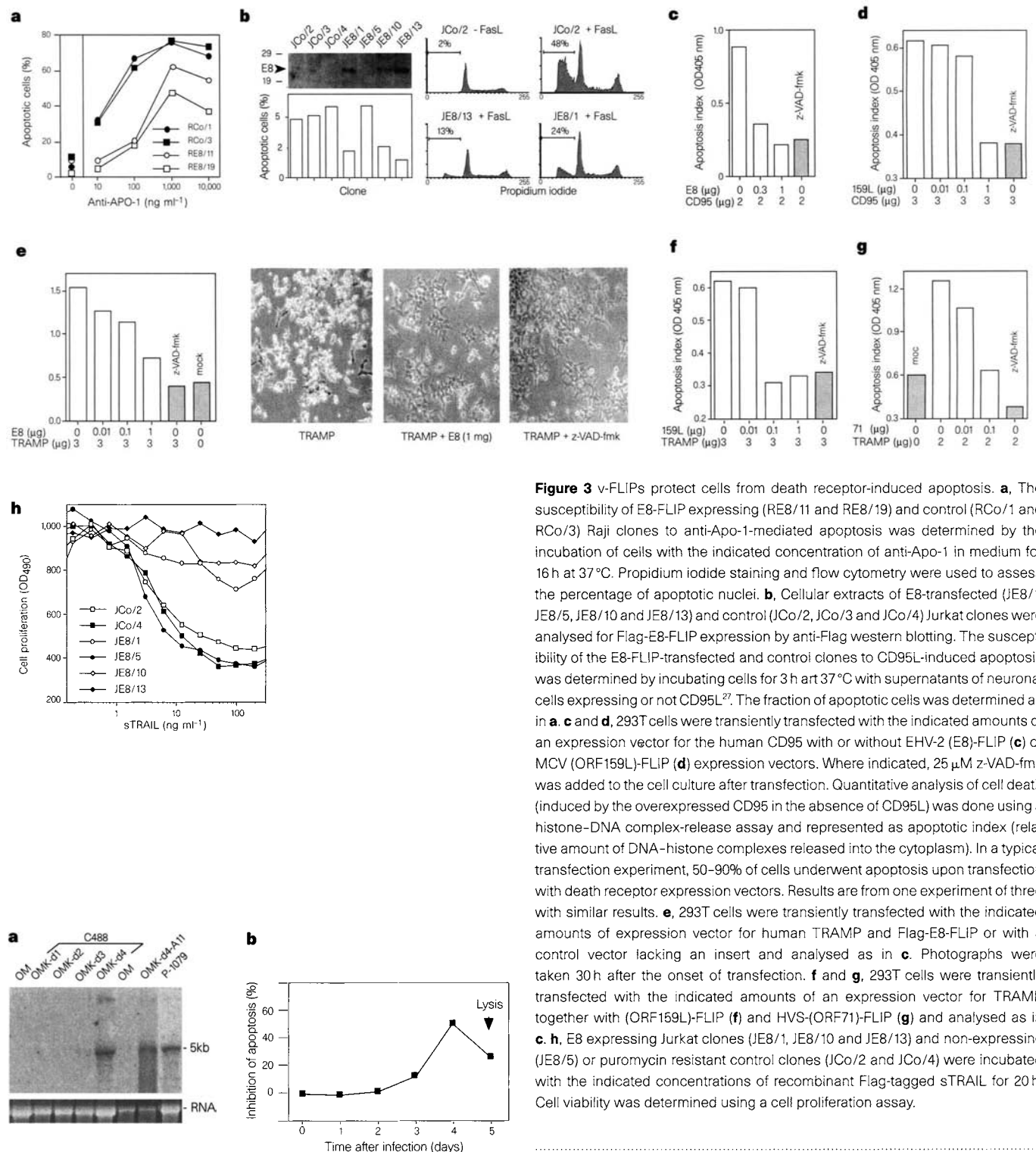


Figure 3 v-FLIPs protect cells from death receptor-induced apoptosis. **a**, The susceptibility of E8-FLIP expressing (RE8/11 and RE8/19) and control (RCO/1 and RCO/3) Raji clones to anti-Apo-1-mediated apoptosis was determined by the incubation of cells with the indicated concentration of anti-Apo-1 in medium for 16 h at 37°C. Propidium iodide staining and flow cytometry were used to assess the percentage of apoptotic nuclei. **b**, Cellular extracts of E8-transfected (JE8/1, JE8/5, JE8/10 and JE8/13) and control (JCo/2, JCo/3 and JCo/4) Jurkat clones were analysed for Flag-E8-FLIP expression by anti-Flag western blotting. The susceptibility of the E8-FLIP-transfected and control clones to CD95L-induced apoptosis was determined by incubating cells for 3 h at 37°C with supernatants of neuronal cells expressing or not CD95L²⁷. The fraction of apoptotic cells was determined as in **a**. **c** and **d**, 293T cells were transiently transfected with the indicated amounts of an expression vector for the human CD95 with or without EHV-2 (E8)-FLIP (**c**) or MCV (ORF159L)-FLIP (**d**) expression vectors. Where indicated, 25 μM z-VAD-fmk was added to the cell culture after transfection. Quantitative analysis of cell death (induced by the overexpressed CD95 in the absence of CD95L) was done using a histone-DNA complex-release assay and represented as apoptotic index (relative amount of DNA-histone complexes released into the cytoplasm). In a typical transfection experiment, 50–90% of cells underwent apoptosis upon transfection with death receptor expression vectors. Results are from one experiment of three with similar results. **e**, 293T cells were transiently transfected with the indicated amounts of expression vector for human TRAMP and Flag-E8-FLIP or with a control vector lacking an insert and analysed as in **c**. Photographs were taken 30 h after the onset of transfection. **f** and **g**, 293T cells were transiently transfected with the indicated amounts of an expression vector for TRAMP together with (ORF159L)-FLIP (**f**) and HVS-(ORF71)-FLIP (**g**) and analysed as in **c**. **h**, E8 expressing Jurkat clones (JE8/1, JE8/10 and JE8/13) and non-expressing (JE8/5) or puromycin resistant control clones (JCo/2 and JCo/4) were incubated with the indicated concentrations of recombinant Flag-tagged sTRAIL for 20 h. Cell viability was determined using a cell proliferation assay.

Figure 4 HVS (ORF71)-FLIP is expressed during the lytic cycle and in semipermissive virus-producing cells. **a**, Northern blot analysis of permissive OMK cells infected with the HVS strain C488 (days indicated refer to the time after infection) or with strain A11 (late conditions of infection). The marmoset monkey-derived semipermissive T cell line (P-1079) produces low titres of HVS particles. **b**, Resistance to CD95L-induced apoptosis in OMK cells expressing HVS-(ORF71)-FLIP. CD95L was added to cells infected with HVS and cell death was compared to non-infected cells. The difference in the susceptibility to CD95L-induced apoptosis (determined by the histone-DNA release assay) is shown. A virus-induced cytopathic effect was evident on day 4 affecting <10% of cells; on day 5 a large proportion (~50%) of the cells not exposed to CD95L was lysed due to viral replication. Results are from one experiment of two with similar results.

Methods

Cell lines and reagents. Human embryonic 293T cells, the human leukaemia Jurkat T-cell line and the human Burkitt lymphoma B-cell line Raji were grown as described⁹. Monoclonal antibodies used in immunoprecipitations and for western blotting included: an anti-Flag antibody and anti-Flag agarose (Kodak International Biotechnologies), an anti-FADD antibody (Transduction Laboratories), and an antibody directed against the Myc epitope (9E10, Sigma). Soluble human TRAIL (amino acids 95–281) was generated by PCR from the EST clone 117926 (Genbank accession number T90422) using oligonucleotides JT403 5'-TCAGCTGCCAGACCTCTGAGGAAAC-3' and JT469 5'-ACTAGTTAGCCAACTAAAAAG-3'; and was cloned into a modified pQE-16 vector (Qiagen) containing the Flag sequence and a linker

(GPGQVQLQ). Protein expression in M15pRep4 bacteria was induced with 0.5 mM IPTG.

Expression vectors. The complete ORF E8 of EHV-2 (viral DNA was a gift from A. J. Davison), ORF 159L of MCV (the plasmid was a gift from G. Darai) and ORF 71 of HVS were amplified by PCR and inserted in frame with an N-terminal Flag epitope into the *EcoRI* site of a vector derived from pCR-3 (Invitrogen). The Flag-E8-FLIP construct was subcloned into the pSRαpuro vector (a gift from R. Sekaly) and then used to create stable puromycin-resistant transfectants of Jurkat and Raji cells. An expression vector for the Myc-tagged cytoplasmic domain of the murine CD95 receptor was generated by insertion of a PCR fragment corresponding to amino acids 166–306 in frame with an N-terminal Myc epitope into a pCR-3 derived vector. The expression vector for human TRAMP and FADD in pCR-3 has been previously described⁹. Human full-length CD95 was subcloned as a *HindIII/XbaI* fragment into pCR-3 and used for transfection of 293T cells.

Cell transfection. Stable puromycin-resistant transfectants of Jurkat and Raji cells were generated by electroporation of 8×10^6 cells in 800 μ l HeBS, mixed with 20 μ g SRαpuro plasmid with or without Flag-E8 insert, at 250 V and 960 μ F. 48 h after transfection, cells were seeded at 10,000 cells per well in flat bottom 96-well plates under selection with 5μ g ml⁻¹ puromycin (Sigma). 293T cells were seeded at $1-2 \times 10^6$ cells per 10-cm plate or $3-6 \times 10^5$ cells per 5-cm plate and transfected the next day by the calcium phosphate precipitation method. The precipitate was left on cells for 8 h, and cells were collected 26–30 h after transfection.

Cell lysis and co-immunoprecipitation. Cell lysis and co-immunoprecipitation of the various tagged proteins were carried out as described⁹. Jurkat and Raji clones or 293T cells were checked for protein expression by anti-tag western blot analysis of postnuclear cell lysates of equivalent protein content. Metabolic labelling of Raji cells with ³⁵S, anti-CD95 immunoprecipitations and 2D-gel electrophoresis were performed as previously described¹⁶. Analysis of DISC-associated FLICE activity was assessed as before¹⁷.

Apoptosis assays. The analysis of apoptosis induced by CD95L was carried out as follows: puromycin-resistant Jurkat clones ($\sim 3 \times 10^5$ cells per 500 μ l) were incubated for 3 h at 37°C with 50 μ l supernatant from neuro-2a cells transfected with a murine CD95L expression vector or with control supernatant from cells transfected with mock vector²⁷. The susceptibility of Raji clones to anti-CD95-induced apoptosis was analysed by incubation of cells (5×10^5 cells ml⁻¹) with varying concentrations of anti-Apo-1 mAb¹⁶ in medium for 16 h at 37°C. Apoptosis was measured by quantifying DNA fragmentation as previously described²⁸. Apoptosis of transiently transfected 293T cells was monitored by the cell-death detection ELISA (Boehringer Mannheim), which detects the presence of soluble histone–DNA complexes. The survival of E8-transfected and control Jurkat clones to TRAIL-induced cell death was tested by incubating cells at 50,000 cells per well in 100 μ l with the indicated concentrations of recombinant TRAIL and 1 μ g ml⁻¹ anti-Flag mAb for 20 h. Proliferating cells were subsequently quantified with the Celltiter 96 AQ proliferation assay (Promega) following the manufacturer's instructions.

Herpesvirus saimiri cultures and transcript analysis. The *in vitro* culture of virus and northern blot analysis of transcripts were carried out as described²². The effect of HVS infection on CD95-mediated cell death of owl monkey kidney (OMK) cells was assayed by seeding the cells at 10^4 cells per well in 96-well flat-bottomed plates. Two days later, half of the wells were infected at a multiplicity of infection of ~ 1 . Recombinant sCD95L (0.3μ g ml⁻¹)⁹ was added together with enhancer (Alexis, San Diego) at different times after infection. Material was collected 20 h later and analysed for the presence of histone–DNA complexes as described.

Received 21 February; accepted 10 March 1997.

- Shen, Y. & Shenk, T. E. Viruses and apoptosis. *Curr. Biol.* **5**, 105–111 (1995).
- Smith, C. A., Farrah, T. & Goodwin, R. G. The TNF receptor superfamily of cellular and viral proteins—activation, costimulation, and death. *Cell* **76**, 959–962 (1994).
- Nagata, S. Apoptosis by death factor. *Cell* **88**, 355–365 (1997).
- Senkevich, T. G. *et al.* Genome sequence of a human tumorigenic poxvirus: prediction of specific host response-evasion genes. *Science* **273**, 813–816 (1996).
- Boldin, M. P. *et al.* A novel protein that interacts with the death domain of Fas/APO1 contains a sequence motif related to the death domain. *J. Biol. Chem.* **270**, 7795–7798 (1995).
- Chinnaiyan, A. M., O'Rourke, K., Tewari, M. & Dixit, V. M. FADD, a novel death domain-containing protein, interacts with the death domain of Fas and initiates apoptosis. *Cell* **81**, 505–512 (1995).

- Muzio, M. *et al.* FLICE, a novel FADD-homologous ICF/CED-3-like protease, is recruited to the CD95 (Fas/APO-1) death-inducing signaling complex. *Cell* **85**, 817–827 (1996).
- Boldin, M. P., Goncharov, T. M., Goltsev, Y. V. & Wallach, D. Involvement of MACH, a novel MORT1/FADD-interacting protease, in Fas/APO-1- and TNF receptor-induced cell death. *Cell* **85**, 803–815 (1996).
- Bodmer, J. L. *et al.* TRAMP, a novel apoptosis-mediating receptor with sequence homology to tumor necrosis factor receptor 1 and Fas/APO-1/CD95. *Immunity* **6**, 79–88 (1997).
- Marsters, S. A. *et al.* Apo-3, a new member of the tumor necrosis factor receptor family, contains a death domain and activates apoptosis and NF- κ B. *Curr. Biol.* **6**, 1669–1676 (1996).
- Kitson, J. *et al.* A death-domain-containing receptor that mediates apoptosis. *Nature* **384**, 372–375 (1996).
- Yu, G. L. *et al.* Signal transduction by DR3, a death domain-containing receptor related to TNFR-1 and CD95. *Science* **274**, 990–992 (1996).
- Chang, Y. *et al.* Identification of herpesvirus-like DNA sequences in AIDS-associated Kaposi's sarcoma. *Science* **266**, 1865–1869 (1994).
- Tschopp, J. & Hofmann, K. Cytotoxic T cells: more weapons for new targets? *Trends Microbiol.* **4**, 91–94 (1996).
- Henkart, P. A. ICE family proteases: mediators of all apoptotic cell death? *Immunity* **4**, 195–201 (1996).
- Kischkel, F. C. *et al.* Cytotoxicity-dependent Apo-1 (Fas/CD95)-associated proteins form a death-inducing signaling complex (DISC) with the receptor. *EMBO J.* **14**, 5579–5588 (1995).
- Medema, J. *et al.* FLICE is activated by association with the CD95 death-inducing signaling complex. *EMBO J.* (in the press).
- Chinnaiyan, A. M. *et al.* FADD/MORT1 is a common mediator of CD95 (Fas/APO-1) and tumor necrosis factor receptor-induced apoptosis. *J. Biol. Chem.* **271**, 4961–4965 (1996).
- Bucher, P., Karplus, K., Moeri, N. & Hofmann, K. A flexible search technique based on generalized profiles. *Comp. Chem.* **20**, 3–24 (1996).
- Fernandes-Alnemri, T. *et al.* *In vitro* activation of CPP32 and Mch3 by Mch4, a novel human apoptotic cysteine protease containing two FADD-like domains. *Proc. Natl Acad. Sci. USA* **93**, 7464–7469 (1996).
- Wiley, S. R. *et al.* Identification and characterization of a new member of the TNF family that induces apoptosis. *Immunity* **3**, 673–682 (1995).
- Fickenscher, H., Biesinger, B., Knappe, A., Wittmann, S. & Fleckenstein, B. Regulation of the herpesvirus saimiri oncogene stpC, similar to that of T-cell activation genes, in growth-transformed human T lymphocytes. *J. Virol.* **70**, 6012–6019 (1996).
- Yang, E. & Korsmeyer, S. J. Molecular thanatopsis: a discourse on the BCL2 family and cell death. *Blood* **88**, 386–401 (1996).
- Huang, D. C. S., Cory, S. & Strasser, A. Bcl-2, Bcl-xL and adenovirus protein E1B19K are functionally equivalent in their ability to inhibit cell death. *Oncogene* **14**, 405–414 (1997).
- Hahne, M. *et al.* Melanoma cells express Fas (Apo-1/CD95) ligand: Implications for tumor immune escape. *Science* **274**, 1363–1366 (1996).
- Strand, S. *et al.* Lymphocyte apoptosis induced by CD95 (Apo-1/Fas) ligand-expressing tumor cells—A mechanism of immune evasion? *Nature Med.* **2**, 1361–1366 (1996).
- Rensing-Ehl, A. *et al.* Local Fas/APO-1 (CD95) ligand-mediated tumor cell killing *in vivo*. *Eur. J. Immunol.* **25**, 2253–2258 (1995).
- Nicoletti, I., Migliorati, G., Pagliacci, M. C., Grignani, F. & Riccardi, C. A rapid and simple method for measuring thymocyte apoptosis by propidium iodide staining and flow cytometry. *J. Immunol. Meth.* **139**, 271–279 (1991).
- Russo, J. J. *et al.* Nucleotide sequence of the Kaposi sarcoma-associated herpesvirus (HHV8). *Proc. Natl Acad. Sci. USA* **93**, 14862–14867 (1996).
- Bertin, J. *et al.* Death effector domain-containing herpesvirus and poxvirus proteins inhibit Fas- and TNFR1-induced apoptosis. *Proc. Natl Acad. Sci. USA* **94**, 1172–1176 (1997).

Acknowledgements. This work was supported by grants from the Swiss National Science Foundation (to J.T.), the European Molecular Biology Organization (to M.T.), the Deutsche Forschungsgemeinschaft (to M.E.P.), the Bundesministerium für Forschung und Technologie, the Tumor Center Heidelberg/Mannheim. H.F. is supported by the Bayerische Forschungsförderung, F.N. by the Mildred Scheel Stiftung and E.M. by the Deutsche Forschungsgemeinschaft. We thank B. Fleckenstein for discussion, M. Irmeler for suggesting the term FLIP and S. Belli for critically reviewing the manuscript.

Correspondence and requests for materials should be addressed to J.T. (e-mail: jurg.tschopp@ib.unil.ch.)

erratum

Ras signalling linked to the cell-cycle machinery by the retinoblastoma protein

Daniel S. Peeper, Todd M. Upton, Mohamed H. Ladha, Elizabeth Neuman, Juan Zalvide, René Bernards, James A. DeCaprio & Mark E. Ewen

Nature **386**, 177–181 (1997).

Due to an error in the reproduction process, the labelling in the right-hand panel of Fig. 5b of this letter was incorrect. The cell type involved is not $Rb^{+/+}$ 3T3, but $Rb^{-/-}$ 3T3, as indicated in the text and the figure legend.

Notes for contributors

Nature is a weekly international journal covering all the sciences. Its readership is interdisciplinary, so manuscripts should be written clearly and simply, authors remembering that English is not the first language of many readers. Competition for *Nature*'s limited space is severe, so brevity is essential. *Nature*'s main formats for original research are Articles and Letters to *Nature*. Lengths of these contributions are specified in terms of numbers of pages they take up. An uninterrupted page of text contains about 1,300 words, so a typical Letter to *Nature*, occupying 2 1/2 pages, can contain about 1,500 words of text and four modest-sized display items (figures and/or tables) with brief legends. There are no formal limits on the number of display items, but extra items must be accompanied by commensurate reductions in the length of the text. Length limits can, however, be varied at the editors' discretion.

Articles are reports of original scientific research whose conclusions represent a substantial advance in understanding of an important problem and are of broad general interest. They do not generally exceed 5 pages of *Nature* and have no more than 50 references. They have a 'heading' of up to 80 words, often rewritten by *Nature*'s editors, which advertises the content of the paper in general terms and which should not contain numbers, abbreviations or measurements unless central to the message of the paper. The first three or four paragraphs of the main text should introduce and summarize the findings. Articles have a few subheadings of not more than 45 characters each.

Letters to *Nature* are short reports of original research focused on an outstanding finding with implications of interest to nonspecialists. They should not generally exceed 2 1/2 pages of *Nature*. They have a bold first paragraph of fewer than 180 words summarizing the background, rationale, conclusion and main implication of the work, aimed particularly at nonspecialist readers. Letters should have no more than 30 references.

Preparation of Articles and Letters

All manuscripts should be in English, typed double-spaced on one side of the paper only. Manuscripts (and revised manuscripts) should not be submitted in electronic form.

Manuscripts should be accompanied by two copies of a brief covering letter from the corresponding author, containing a full postal address, telephone and fax numbers, and e-mail address. Articles and Letters contain a statement at the end of the text: "Correspondence and requests for materials to xxxxx", with an e-mail address if desired, followed by the accession numbers or Web site address for supporting datasets. The designated author need not be the senior author, but *Nature* will expect this author to handle matters arising from the published contribution.

Five copies of manuscripts (and revised manuscripts) and figures are required. Each copy of a revised manuscript should be accompanied by a copy of the author's response to referees' and editors' comments on the earlier version. Two copies of relevant papers by any of the authors submitted or in press elsewhere (including papers submitted elsewhere while the *Nature* manuscript is under consideration) should also be enclosed, clearly marked as such; failure to do so may prejudice consideration of the manuscript. Titles should not exceed three lines of 28 characters each (including spaces), and should avoid numerical values, acronyms, abbreviations and punctuation wherever possible.

References should be numbered sequentially as they appear in the text, tables and figure legends. Titles of all cited articles should be provided. Only papers that have been published or submitted to a named journal should be in the reference list: manuscripts in preparation and personal communications should be mentioned in the text with a list of authors. Published conference abstracts may be included in the reference list, but text, footnotes, grant details and acknowledgements may not.

Acknowledgements must be brief and appear after the reference list. *Nature* does not publish grant or contribution numbers.

Tables should each be presented on a separate sheet of paper. Symbols and abbreviations should be defined immediately below the table, fol-

lowed by essential descriptive material, all in double-spaced text.

Figures should not be larger than 22 by 28 cm unless unavoidable, and should not contain more than one panel unless the assembled parts are logically connected. Each copy should be marked with the corresponding author's name and, when known, the manuscript reference number. One photocopy of original figures should be provided. All original figures are returned when *Nature* cannot offer to publish a manuscript, but one copy of the manuscript and the photocopy of the figures are kept in confidential files. Colour figures are welcome, but a contribution towards the cost of reproduction is requested. Inability to pay this charge will not prevent *Nature* from publishing figures for which colour is essential. Once a paper is accepted for publication, electronic versions of figures are preferred.

Methods. Brief details of materials and methods can be given in the text. Otherwise, they should be described at the end of the text in a section entitled "Methods", subdivided by short, bold headings referring to the methods used. This section should not exceed 800 words and ideally should be shorter. In contributions with methods sections, figure legends should not exceed 100 words each, and should be confined to a description of the figure and the symbols used.

Materials. As a condition of publication, authors are required to make materials and methods used freely available to academic researchers for their own use. Supporting datasets must be made available by the authors both directly and by posting on *Nature*'s Web site, by deposition in the appropriate public database or by distribution on the Internet.

Submissions. Manuscripts can be submitted to the Editor at *Nature*, Porters South, 4 Crinan St, London N1 9XW or at *Nature*, 968 National Press Building, Washington, DC 20045-1938. Proofs should be returned by express mail to London. Fax numbers and e-mail addresses for all editorial offices are on *Nature*'s information page.

Presubmission enquiries for Articles and Letters should be made by e-mail to nature@nature.com, clearly labelled as such. They should include a fully referenced summary paragraph in *Nature* format, and be accompanied by a paragraph stating briefly why the submission will be of interest to a broad general scientific readership. Authors will usually receive a response within two working days of receipt.

Material submitted to *Nature* must not be discussed with the media. Accepted contributions may be discussed with the media a week before the publication date under *Nature*'s embargo conditions. (See *Nature* 383, 745; 1996 for further explanation.)

Other contributions

Nature publishes other contributed material as follows:

Correspondence contributions are short comments on topical issues, anecdotal material or reactions of readers to non-technical material published in *Nature*. All are unsolicited; submissions by e-mail to corres@nature.com are preferred (accents should be clearly indicated).

Commentary articles are informal editorials on topical issues of public interest that have some bearing on scientific research. They are often controversial. Unsolicited manuscripts are considered, but proposals can be made to the Commentary Editor by e-mail (nature@nature.com).

News and Views articles inform nonspecialist readers about new scientific advances, as reported either in published papers or at scientific meetings. Most are commissioned, but proposals can be made to the News and Views Editor before the paper is published or the meeting takes place. Detailed guidelines are available on request.

Scientific Correspondence contributions are brief reports of general interest, and are usually peer-reviewed. Priority is given to contributions of fewer than 500 words. Detailed guidelines are available on request. Presubmission enquiries are not considered: contributions for Scientific Correspondence should be submitted by mail, not e-mail.

Review Articles survey recent developments in a topical area of research. They should not generally occupy more than 6 pages of *Nature*.

Progress articles are shorter reviews of particularly topical and fast-moving fields. They do not exceed 4 pages of *Nature*.

Contributors wishing to submit a Review Article or Progress should first send an enquiry or a brief synopsis to the Reviews Coordinator, from whom guidelines are available on request.

Experimental biology

Here are a few items that may be exhibited at the annual FASEB meeting being held in New Orleans this month – a new expression detection system, a CO₂ incubator, a protein sequencing system and conductive pipettes.

HP 1100 Series

From Hewlett-Packard

Complementary N- and C-terminal sequence information in one protein-sequencing system

This system is stated to use highly efficient column-based, amino-terminal sequencing chemistry, as well as a practical approach for automated carboxy-terminal sequencing. The new system is fully integrated with the HP 1100 Series HPLC and diode-array detector. The identification of both N-terminal and C-terminal amino-acid sequences should give bioscientists the information needed to unambiguously determine and verify the sequences of both native and recombinant protein products from end to end. Using this system, it is possible to detect sample heterogeneity, such as the families of 'ragged' C- or N-terminal proteins that are the result of post-translational protein processing. N-terminally blocked protein samples and proteins with amino-acid truncations, deletions and substitutions are also said to be analysed easily. The chemical efficiency of this Edman-based, N-terminal sequencing chemistry is said to result in quick cycles that produce sequence information well beyond 80 cycles for both liquid and electroblotted samples. A high repetitive yield for picomole amounts of sample applied to the sequencing support extends the available sequence information. Automated C-terminal protein-sequencing technology replaces conventional manual methods, such as carboxypeptidase treatments, which can often yield variable and limited results.

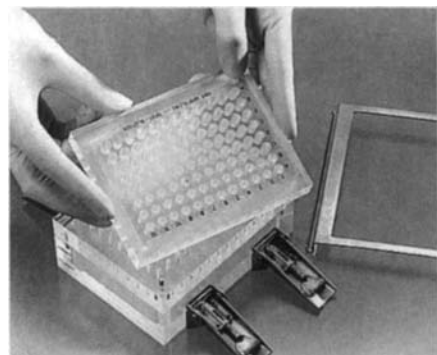
Reader Enquiry No. 100

Minifold Slot-Blot unit

From Schleicher & Schuell

A new unit for dot- and slot-blotting

With this product, sample volumes are concentrated as small slots on an area of



Minifold can screen 96 samples at time

only 2 mm². This is said to result in high signal intensity, ensuring very high sensitivity that accompanies the reproducible densitometry quantification of the slot format. Using the Minifold Spot-Blot unit, 96 samples can be screened simultaneously in convenient microtitre plate format. The metal clamping frame and the adjustable clamps have simplified assembling.

Reader Enquiry No. 101

Interleukin 12 ELISA kits

From Genzyme

Four new interleukin 12 (IL-12) ELISA kits, including mouse total IL-12, mouse IL-12 p70, human total IL-12 and human IL-12 p70

Separate measurement of p70 and total IL-12 can be useful in determining the physiological role of IL-12 in several research models. IL-12 is said to play an important role in the regulation of cellular immune responses and is therefore a useful tool in investigating Th1/Th2. Specifically, researchers have found that IL-12 regulates organ-specific and systemic autoimmunity processes, infectious diseases and shows anti-tumour activity in mice. The new IL-12 ELISAs complement a full line of mouse and human cytokine kits useful in several applications for the investigation of Th1/Th2 and pro-inflammatory cytokine research. The kits are stated to be sensitive, specific and easy to use. They are also available as single kits or in the economical multiplate bulk kit format.

Reader Enquiry No. 102

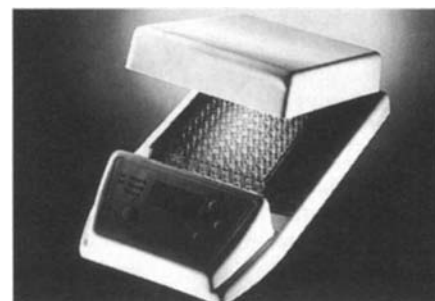
PINK expression system

From R&D Systems Europe

Part of the Ingenius range, the system allows production of visible fusion protein in Escherichia coli

The pINK vectors encode an N-terminal tag, cytochrome b5, which confers a pink colour on fusion proteins. This colour makes it easy to track the protein of interest during isolation and purification. In addition, a specific spectrophotometric measurement can be easily performed to quantify the amount of protein produced. The vectors contain the *phoA* promoter for inducible expression in *E. coli*. A His tag is available for simple purification of fusion proteins and a factor Xa site allows removal of the tags from the protein of interest. Three reading frame versions are available.

Reader Enquiry No. 103



The Solo personal incubation unit from Denley

Incubation

Solo incubator

From Denley Instruments

New single-plate incubator for personal use

Microplate users should no longer have to plan assays round the schedule of a large capacity incubator with the advent of this new single-plate incubator. The incubator itself is little bigger than a microplate and operates up to a temperature of 60 °C with an accuracy of ± 5 per cent. Equally important is well-to-well temperature variation, which is said to be less than one per cent. Solo also successfully addresses the problem of evaporative loss during long incubation periods. According to the manufacturer, no more than 1.5 μ l per well is lost at 37 °C from the incubated plates. With a low-voltage d.c. power requirement, it can easily moved from one location to another. The unit is operated via a simple keypad and clear LED display. Worth noting is the special 'set time' feature, which allows the user to set a count-down period of up to 180 min in 5-minute increments. The 'set temperature' can also be changed at any time and the incubator will adjust immediately. Any difference between the set temperature and current temperature will be indicated by both visual and audio alarms when the time is started.

Reader Enquiry No. 104

BBD 6220 incubator

From Heraeus

This new CO₂ incubator offers in situ sterilization

BBD 6220 incubator claims to be the only CO₂ incubator with an in-built 180 °C automatic hot-air sterilization cycle. Specifically developed to offer increased protection against contamination, the new incubator features a fully automatic 180 °C hot-air disinfection cycle, which destroys all viruses, bacteria and fungi commonly encountered in cell culture applications.

The disinfection process can be carried out with sensors and all internal fittings in position, avoiding the further risk of contamination often associated with removal and refitting of such components. Additional protection is provided by a pyrolytic germ barrier, which sterilizes the water vapour used for humidification before it reaches the working chamber. There is no water storage within the incubator chamber — water is stored externally in a reservoir which is said to be easy to fill and empty. The BBD 6220 controls temperature, CO₂ level and relative humidity with the precision necessary to recreate the environmental conditions required for successful *in vitro* growth of cell cultures. The hot-air, jacketed heating system prevents condensation forming on the stainless steel interior, even at a relative humidity of 95 per cent. Over-temperature protection and an error diagnosis system supplement the safety features of the incubator. Following a power failure, any incubation settings are automatically restored. All operating parameters can be monitored, controlled and documented via an RS 232 digital interface, which is fitted as standard.

Reader Enquiry No. 105

Consumables

Nanosep MF

From Pall Filtron

A microfiltration device that uses Pall nylon Bio-Inert membranes

This centrifugal concentrator is made of modified nylon, which produces a membrane that provides yields stated to be greater than 90 per cent. The concentrator has a large 500- μ l sample reservoir, allowing for the processing of large gel slices. The device is suitable for separating DNA from agarose gels, separating proteins, oligonucleotides and RNA from acrylamide gels, as well as clarifying samples before HPLC analysis. It is said to be easy to use and to have less hold-up than a syringe filter.



Pall Filtration's Nanosep MF for centrifugation

Moreover, it is compatible with a wide variety of organic solvents and chemicals.

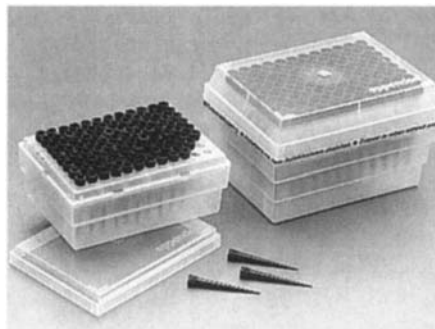
Reader Enquiry No. 106

Conductive pipette tips

From Eppendorf

For use with automatic dispensing, these tips recognize every filling height and need to be only minimally immersed to pipette

According to the manufacturer, this tip has been created to eliminate the possibility of dispensing errors. Contamination-free work is ensured by discarding the tips after dispensing. A further advantage of the conductive pipette tip is that solutions can be taken from, and dispensed into, a wide range of different containers. For example, it is possible to work



Eppendorf's conductive pipettes recognize filling height in automated dispensing systems

simultaneously with sample tubes, microtitre plates and test tubes. Conductive pipette tips are available in 200- μ l and 1,000- μ l sizes. They are manufactured in boxes of 96 and are, of course, autoclavable.

Reader Enquiry No. 107

These notes are compiled by Brendan Horton from information provided by the manufacturers. For more details, fill in the reader service card bound inside the journal.

NEWCASTLE UNIVERSITY
MOLECULAR BIOLOGY UNIT



PROTEIN SEQUENCING

Membrane-bound, solid or solution samples

£12 per cycle

5 cycles minimum - no set up charge

For further information contact Dr. Joe Gray

Phone: +44 191 222 8612

Fax: +44 191 222 8129

Email: JOE.GRAY@newcastle.ac.uk

READER ENQUIRY NO. 69



PEPTIDE TECHNOLOGIES

- Custom Peptide Synthesis
- Purified and Analyzed for Your Research
- Biotinylated, Dye-labelled, Phosphopeptides, etc.
- Peptide Vaccines: KLH Conjugates and MAPS
- CALL 301 869 7306 (USA)
- FAX 301 869 7308
- URL <http://www.luxnet.com/peptide/>

READER ENQUIRY NO. 89

DIABORM® EQUILIBRIUM DIALYSIS SYSTEM



The gentlest method to quantify the binding parameters and thermodynamic data between ligands and biopolymers. Results come closest to the actual values. Sample volume from 0.2 ml to 5.0 ml. Up to 20 experiments in parallel.

DIABORM
Gentle
P.O. Box 550124
D 81215 Munich

READER ENQUIRY NO. 61

ADVERTISEMENTS

Gramsch Laboratories

Peptide Synthesis (25 μ moles):

10 AA DM 650.00; 15 AA DM 730.00;

20 AA DM 780.00

buy only when satisfied with the HPLC profile

Antiserum Production Service:

Package : DM 2100.00 (Peptide/Conjug./Serum)

New Antisera:

Hum, Mouse, Rat Somatost. Rec. 1/2A/2B/3/4/5

Nociceptin; G-Protein subunits α , β , γ ;

Distributor for Japan: FUNAKOSHI CO, Tokio 113

Gramsch Laboratories

D-85247 SCHWABHAUSEN, Germany

Phone +49-8138-1769 or +49-8131-5679-0

Fax +49-8138-1260 or +49-8131-5679-24

READER ENQUIRY NO. 94

HIGH QUALITY PEPTIDE SYNTHESIS AND PROTEIN SEQUENCING

Fast delivery

Competitive prices

Also advisory services

Contact:
Dr J A Smith
Chairman

Pepsyn Limited, The University of Liverpool
Life Science Building, Crown Street,
Liverpool L69 7ZB

Tel: 0151 794 4329 Fax: 0151 794 4349

E-mail: J.A.Smith@liverpool.ac.uk

READER ENQUIRY NO. 93